

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISH	IED U	UNDER THE PATENT COOPERATION TREATY (PCT)
(51) International Patent Classification 6:		(11) International Publication Number: WO 99/09164
C12N 15/12, C07K 14/47, C12Q 1/68, A61K 48/00, C07K 16/18, A61K 38/17	A1	(43) International Publication Date: 25 February 1999 (25.02.99)
(21) International Application Number: PCT/US	98/169	05 (74) Agent: HALLUIN, Albert, P.; Howrey & Simon, Box 34, 1299 Pennsylvania Avenue N.W., Washington, DC 20004 (US).
(22) International Filing Date: 14 August 1998 (1	14.08.9	98)
(30) Priority Data: 60/055,784 15 August 1997 (15.08.97) 60/064,926 7 November 1997 (07.11.97) 60/065,367 12 November 1997 (12.11.97) 09/071,715 1 May 1998 (01.05.98) 09/084,471 22 May 1998 (22.05.98) (71) Applicant: ONCORMED, INC. [US/US]; 205 Perry 19 Gaithersburg, MD 20877 (US).) (7) (1	(81) Designated States: AU, CA, IL, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

(72) Inventors: MURPHY, Patricia, D.; 16 Stockbridge Road, Slingerlands, NY 12159-0639 (US). WHITE, Marga, B.; 8323 Sharon Drive, Frederick, MD 21704 (US). RABIN, Mark, B.; 1516 Defoe Street, Rockville, MD 20850 (US). OLSON, Sheri, J.; 2854 Yarling Court, Falls Church, VA 22042 (US). YOSHIKAWA, Matthew; 19300 Circle Gate Drive, Germantown, MD 20874 (US). JACKSON, Geoffrey, M.; 3312 Major Denton Drive, Beltsville, MD 20705 (US). ESKANDARI, Tara; 5440 Marinelli Road, Rockville, MD 20852 (US). SCHRYER, Brenda; 4240 S. Harbor Boulevard, Oxnard, CA 93035 (US). PARK, Michael; 770 Ivy League Lane, Rockville, MD 20850 (US).

(57) Abstract

Five DNA and protein sequences have been determined for the BRCA2 gene, as have been ten polymorphic sites and their rates of occurrence in the normal alleles of BRCA2. The sequences BRCA2(ormi1-5) and the ten polymorphic sites will provide accuracy and reliability for genetic testing. One skilled in the art will be able to avoid misinterpretations of changes in the gene and/or protein sequence, determine the presence of a normal sequence, and of mutations of BRCA2. This invention is also related to a method of performing gene therapy with BRCA2(omi1-5) coding sequences or fragments thereof. This invention is further related to protein therapy with BRCA2(omi1-5) proteins or their functional equivalents.

> ATTORNEY DOCKET NUMBER: 9301-123 SERIAL NUMBER: 09/724,538 REFERENCE: CY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
		GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BE BF	Belgium Burkina Faso	GR	Greece	.****	Republic of Macedonia	TR	Turkey
		HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BG	Bulgaria	IE	Ireland	MN	Mongolia	UA	Ukraine
BJ	Benin	_	Israel ·	MR	Mauritania	UG	Uganda
BR	Brazil	IL	Iceland	MW	Malawi	US	United States of Americ
BY	Belarus	IS		*	Mexico	UZ	Uzbekistan
CA	Canada	IT	Italy	MX	•	VN	Viet Nam
CF	Central African Republic	JP	Japan	NE	Niger	YU	
CG	Congo	KE	Kenya	NL	Netherlands		Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
KE	Estonia	LR	Liberia	SG	Singapore		

CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

This is an U.S. utility patent application based on U.S. Provisional Application Serial Nos. 60/055,784 filed on August 15, 1997, 60/064,926 filed on November 7, 1997, and 60/065,367 filed on November 12, 1997.

FIELD OF THE INVENTION

This invention relates to a gene which has been associated with breast cancer where the gene is found to be mutated. More specifically, this invention relates to five unique coding sequences of BRCA2 gene BRCA2^(omi1), BRCA2^(omi2), BRCA2^(omi3), BRCA2^(omi4), and BRCA2^(omi5) identified in human subjects which define five novel haplotypes.

BACKGROUND OF THE INVENTION

It has been estimated that about 5-10% of breast cancer is inherited (Rowell, S., et al., American Journal of Human Genetics 55:861-865 (1994)). The first gene associated with both breast and ovarian cancer was cloned in 1994 from chromosome 17 by Miki, Y., et al., Science 266:66-71 (1994). A second high-risk breast cancer conferring gene was located on chromosome 13 in 1994 (Wooster, R., et al., Science 265:2088-2090) and subsequently cloned in 1995 (Wooster, R., et al., Nature 378:789-792). Mutations in this "tumor suppressor" gene are thought to account for roughly 35% of inherited breast cancer and 80-90% of families with male breast cancer.

Locating one or more mutations in the BRCA2 region of chromosome 13 provides a promising approach to reducing the high incidence and mortality associated with breast cancer through the early detection of women and men at high risk. These individuals, once identified, can be targeted for more aggressive prevention programs. Screening is carried out by a variety of methods which include karyotyping, probe binding and DNA sequencing.

In DNA sequencing technology, genomic DNA is extracted from whole blood and the coding regions of the BRCA2 gene are amplified. Each of the coding regions may be sequenced completely and the results are compared to the normal DNA sequence of the gene. Alternatively, the coding sequence of the sample gene may be compared to a panel of known mutations or other

25

30

5

10

15

20

5

10

15

20

25

30

screening procedure before completely sequencing the gene and comparing it to a normal sequence of the gene.

The BRCA2 gene is divided into 27 separate exons. Exon 1 is noncoding, in that it is not part of the final functional BRCA2 protein product. The BRCA2 coding region spans roughly 10433 base pairs (bp) over 70 kb. Each exon consists of 100-600 bp, except for exons 10, 11 and 27. The full length mRNA is 11-12 kb. To sequence the coding region of the BRCA2 gene, each exon is amplified separately and the resulting PCR products are sequenced in the forward and reverse directions. Because exons 10, 11, and 27 are so large, we have divided them into three, twenty-one, and two overlapping PCR fragments (respectively) of approximately 250-625 bp each (segments "A" through "C" of exon 10, "A" through "U" of exon 11, and "A" through "B" of exon 27).

Many mutations and normal polymorphisms have already been reported in the BRCA2 gene. A world wide web site has been built to facilitate the detection and characterization of alterations in breast cancer susceptibility genes. Such mutations in BRCA2 can be accessed through the Breast Cancer Information Core (BIC) at http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995). The information on BRCA2 was added in February, 1996.

The genetics of Breast Cancer Syndrome is autosomal dominant with reduced penetrance. In simple terms, this means that the syndrome runs through families: (1) both sexes can be carriers (mostly women get the disease but men can both pass it on and occasionally get breast cancer); (2) most generations will likely have breast cancer; (3) occasionally women carriers either die young before they have the time to manifest disease (and yet have offspring who get it) or they never develop breast or ovarian cancer and die of old age (the latter people are said to have "reduced penetrance" because they never develop cancer). Pedigree analysis and genetic counseling is absolutely essential to the proper workup of a family prior to any lab work.

Until now, the only sources of genomic sequence information for BRCA2 were GenBank (Accession Number U43746), or through the Breast Information Core (BIC) database on the Internet which requires membership in the BIC consortium. However, based upon the disclosure of this patent application, in neither GenBank nor BIC were the sequences identified and listed entirely accurate. There is a need in the art to correct these mistakes which otherwise may lead to misinterpretation of the sequence data from the patient as abnormal when it was not, or vice versa.

In addition, there is a need in the art to have available a functional allele profile which represents the most likely BRCA2 sequences to be found in the majority of the normal population. This functional allele profile is based upon frequent polymorphisms and the correct backbone sequence. The knowledge of several common normal haplotypes will make it possible for true mutations to be easily identified or differentiated from polymorphisms. Identification of mutations of the BRCA2 gene and protein would allow more widespread diagnostic screening for hereditary breast cancer than is currently possible.

The use of these common normal haplotypes, in addition to the previously published BRCA2 sequence, will reduce the likelihood of misinterpreting a "sequence variation" found in the normal population with a pathologic "mutation" (i.e. causes disease in the individual or puts the individual at a high risk of developing the disease). With large interest in breast cancer predisposition testing, misinterpretation is particularly worrisome. People who already have breast cancer are asking the clinical question: "is my disease caused by a heritable genetic mutation?" The relatives of the those with breast cancer are asking the question: "Am I also a carrier of the mutation my relative has? Thus, is my risk increased, and should I undergo a more aggressive surveillance program?"

SUMMARY OF THE INVENTION

5

10

15

20

25

30

The present invention is based on the discovery of the correct genomic BRCA2 sequence and five novel sequence haplotypes found in normal human subjects of the BRCA2 gene.

5

10

15

20

25

30

It is an object of this invention to provide the correct intronic/exonic sequence of the BRCA2 gene.

It is another object of this invention to provide five unique haplotype sequences of the BRCA2 gene in normal individuals which do not correspond to increased cancer susceptibility.

It is another object of this invention to sequence a BRCA2 gene or a portion thereof and compare it to the five haplotype sequences to determine whether a sequence variation noted represents a polymorphism or a potentially harmful mutation.

It is another object of this invention to provide a list of the pairs which occur at each of ten polymorphic points in the BRCA2 gene.

It is another object of this invention to provide the rates of occurrence for the polymorphisms at codons 289, 372, 455, 743, 894, 991, 1132, 1269, 2414, and 2951 in the BRCA2 gene.

It is another object of this invention to provide a method wherein all exons of BRCA2 gene or parts thereof, are amplified with one or more oligonucleotide primers.

It is another object of this invention to provide a method of identifying a individual who carries no mutation(s) of the BRCA2 gene and is therefore at no increased risk or susceptibility to breast or ovarian cancer based on a finding that the individual does not carry an abnormal BRCA2 genes.

It is another object of this invention to provide a method of identifying a mutation in BRCA2 gene leading to predisposition or higher susceptibility to breast or ovarian cancer.

It is another object of this invention to provide five novel BRCA2 protein sequences derived from five BRCA2 haplotype sequences.

It is another object of the invention to encompass prokaryotic or eukaryotic host cells comprising an expression vector having a DNA sequence that encodes for all or a fragment of the five novel BRCA2 protein sequences, a BRCA2 polypeptide thereof, or a functional equivalent thereof.

It is another object of the invention to encompass an anti-BRCA2 protein antibody using all of fragments of the five novel BRCA2 protein

sequences, a BRCA2 polypeptide thereof or a functional equivalent thereof as an immunogen.

There is a need in the art for cDNA sequences of the BRCA2 gene and for the protein sequences of BRCA2 gene from normal individuals who are not at risk for increased susceptibility for cancer. In order to determine whether a sample from a patient suspected of containing a BRCA2 mutation actually has the mutation, the patient's BRCA2 DNA and/or amino acid sequence need to be compared to all known normal BRCA2 sequences. Failure to compare the sequence obtained to all naturally occurring normal sequences may result in reporting a sample as containing a potentially harmful mutation when it is a polymorphism without clinical significance.

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- identifying individuals having a normal BRCA2 gene with no coding sequence mutations, who therefore cannot be said to have an increased genetic susceptibility to breast or ovarian cancer from their BRCA2 genes;
- avoiding misinterpretation of normal polymorphisms found in the BRCA2 gene;
- c) determining the presence of a previously unknown mutation in the BRCA2 gene;
 - identifying a mutation in exon 11 of BRCA2 which indicates a predisposition or higher susceptibility to ovarian cancer than

breast

25

5

10

15

- cancer (i.e., resides in the putative "ovarian cancer cluster" region);
 - e) probing a human sample of the BRCA2 gene by allele to determine the presence or absence of either polymorphic alleles or mutations;
- of) performing gene therapy with the correct BRCA2 gene sequence.
 - g) performing protein replacement therapy with the correct BRCA 2 protein sequence or a functional equivalent thereof.

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 shows the GenBank genomic sequence of BRCA2 (Accession Number U43746). The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Incorrect exonic sequences at exons 5 and 16 are shown with boldface type.

FIGURE 2 shows the corrected genomic sequence of BRCA2. The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Corrected intronic and exonic sequences at exons 5, 11 and 15 are shown with boldface type.

FIGURE 3 shows the alternative alleles at polymorphic sites along a chromosome which can be represented as a unit or "haplotype" within a gene such as BRCA2. The haplotype that is in GenBank (GB) is shown with light shading. Five additional haplotypes are shown in FIGURE 3

(encompassing the alternative alleles found at nucleotide sites 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470 and 9079). BRCA2 (omi-1), BRCA2 (omi-2), BRCA2 (omi-3), BRCA2 (omi-4), and BRCA2 (omi-5) are represented with mixed dark and light shading (numbers 2, 4, 6, 8 and 10 from left to right). In total, 5 of 10 haplotypes along the BRCA2 gene are unique.

20

25

30

15

5

10

DETAILED DESCRIPTION OF THE INVENTION

DEFINITIONS

The following definitions are provided for the purpose of understanding this invention.

"Breast and Ovarian cancer" is understood by those skilled in the art to include breast, ovarian and pancreatic cancer in women and also breast, prostate and pancreatic cancer in men. BRCA2 is associated with genetic susceptibility to breast, ovarian and pancreatic cancer. Therefore, claims in this document which recite breast and/or ovarian cancer refer to breast, ovarian, prostate, and pancreatic cancers in men and women.

"Coding sequence" refers to those portions of a gene which, taken together, code for a peptide (protein), or which nucleic acid itself has function.

"Protein" or "peptide" refers to a sequence of amino acids which has function.

"BRCA2^(omi)" refers to the genomic BRCA2 sequence disclosed in Genbank (Accession Number U43746) wherein,

5

10

15

20

25

30

(1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5; and

(2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon.

"BRCA2^(omi 1-5)" refers to five unique DNA sequences of the BRCA2 gene and their introns (particularly the slice sites adjacent to the exons). These sequences were found by end to end sequencing of the BRCA2 gene from 5 individuals randomly drawn from the population and who were documented to have no family history of breast or ovarian cancer. The sequenced exons were found not to contain any truncating mutations. In all cases the change of a nucleic acid at a polymorphic site lead to a codon change and a change of amino acid from the previously published standard in GenBank (see TABLE III). In some cases the frequency of occurrence of a nucleic acid change was found to differ from the published frequency or was newly determined. These sequence variations are believed to be alleles whose haplotypes do not indicate an increased risk for cancer.

"Normal DNA sequence" also called "normal gene sequence" refere to a nucleic acid sequence, the nucleic acid of which are known to occur at their respective positions with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which itself has normal function.

"Normal Protein Sequence" refers to the protein sequence, the amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein.

"Normal Sequence" refers to the nucleic acid or protein sequence, the nucleic or amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which nucleic acid itself has a normal function.

"Haplotype" refers to a series of specific alleles within a gene along a chromosome.

"Functional allele profile" refers a list of those alleles in the normal population which have the funll function.

5

10

15

20

25

30

"Mutation" refers to a base change or a gain or loss of base pair(s) in a DNA sequence, which results in a DNA sequence coding for a non-functional protein or a protein with substantially reduced or altered function.

"Polymorphism" refers to a base change in a DNA sequence which is not associated with known pathology.

"Primer" refers to a sequence comprising about 15 or more nucleotides having a sequence complementary to the BRCA2 gene. Other primers which can be used for primer hybridization will be known or readily ascertainable to those skilled in the art.

"Substantially complementary to" refers to primer sequences which hybridize to the sequences provided under stringent conditions and/or sequences having sufficient homology with BRCA2 sequences, such that the allele specific oligonucleotide primers hybridize to the BRCA2 sequences to which they are complimentary.

"Isolated nucleic acids" refers to nucleic acids substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which they may be associated. Such association is typically either in cellular material or in a synthesis medium.

"Biological sample" or "body sample" refers to a sample containing DNA oatained from a biological source. The sample may be from a living, dead or even archeological source from a variety of tissues and cells. Examples include body fluid (e.g. blood (ieukocytes), urine (epithelial cells), saliva, breast milk, menstrual flow, cervical and vaginal secretions, etc.), skin, hair roots/follicle, mucus membrane (e.g. buccal or tongue cell scrapings), cervicovaginal cells (from PAP smear, etc.), lymphatic tissue, internal tissue (normal or tumor).

"Vector" refers to any polynucleotide which is capable of self replication or inducing integration into a self-replicating polynucleotide. Examples include polynucleotides containing an origin or replication or an

integration site. Vectors may be intergrated into the host cell's chromosome or form an autonomously replicating unit.

"A tumor growth inhibitor" refers to a molecule such as, all or a fragment of BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent thereof that is effective for preventing the formation of, reducing, or eliminating a transformed or malignant phenotype of breast or ovarian cancer cells.

"A BRCA2 polypeptide" refers to a BRCA2 polypeptide either directly derived from the BRCA2 protein, or homologous to the BRCA2 protein, or a fusion protein consisting of all or fragments of the BRCA2 protein and polypeptides.

"A functional equivalent" refers to a molecule including an unnatural BRCA2 polypeptide, a drug or a natural product which retains substantial biological activity as the native BRCA2 protein. The activity and function of BRCA2 protein may include transactivation, granin, DNA repair, among others.

"A target polynucleotide" refers to the nucleic acid sequence of interest, for example, the BRCA2 encoding polynucleotide. Other primers which can be used for primer hybridization will be known or readily ascertainable to those of skill in the art.

The invention in several of its embodiments includes: an isolated DNA sequence of the BRCA2 coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, and 12, a protein sequence of the BRCA2 protein as set forth in SEQ ID NO:5, 7, 9, 11, i3, a method of identifying individuals having a normal BRCA2 gene with no increased risk for breast and ovarian cancer, a method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, a method of performing gene therapy to prevent or treat a tumor, a method of protein replacement therapy to prevent or treat a tumor, a diagnostic reagent comprising all or fragments of the disclosed BRCA2 cDNA and protein sequences.

SEQUENCING

5

10

15

20

25

30

5

10

15

20

25

30

Any nucleic acid specimen, in purified or non-purified form, can be utilized as the starting nucleic acid, providing it contains, or is suspected of containing, the specific nucleic acid sequence containing a polymorphic or a mutant allele. Thus, the process may amplify, for example, DNA or RNA, including mRNA and cDNA, wherein DNA or RNA may be single stranded or double stranded. In the event that RNA is to be used as a template, enzymes and/or conditions optimal for reverse transcribing the template to DNA would be utilized. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture of nucleic acids may also be employed, or the nucleic acids produced in a previous method such as an amplification reaction using the same or different primers may be so utilized. The specific nucleic acid sequence to be amplified, i.e., the polymorphic and/or the mutant allele, may be a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. A variety of amplification techniques may be used such as ligating the DNA sample or fragments thereof to a vector capable of replication or incorporation into a replicating system thereby increasing the number of copies of DNA suspected of containing at least a portion of the BRCA2 gene. Amplification techniques include so called "shot gun cloning". It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as contained in whole human DNA.

It should be noted that one need not sequence the entire coding region or even an entire DNA fragment in order to determine whether or not a mutation is present. For example, when a mutation is known in one family member, it is sufficient to determine the sequence at only the mutation site by sequencing or by other mutation detection systems such as ASO when testing other family members.

DNA utilized herein may be extracted from a body sample, such as blood, tissue material and other biological sample by a variety of techniques such as that described by Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, 1982). If the extracted sample is impure, it may be treated before amplification with an amount of a reagent

5

10

15

20

25

30

effective to open the cells, and to expose and/or separate the strand(s) of the nucleic acid(s). This lysing and nucleic acid denaturing step to expose and separate the strands will allow amplification to occur much more readily.

For amplification by cloning, the isolated DNA may be cleaved into fragments by a restriction endonuclease or by shearing by passing the DNA containing mixture through a 25 gauge needle from a syringe to prepare 1-1.5 kb fragments. The fragments are then ligated to a cleaved vector (virus, plasmid, transposon, cosmid etc.) and then the recombinant vector so formed is then replicated in a manner typical for that vector.

For a PCR amplification, the deoxyribonucleotide triphosphates dATP. dCTP, dGTP, and dTTP are added to the synthesis mixture, either separately or together with the primers, in adequate amounts and the resulting solution is heated to about 90°-100°C from about 1 to 10 minutes. preferably from 1 to 4 minutes. After this heating period, the solution is allowed to cool, which is preferable for the primer hybridization. To the cooled mixture is added an appropriate agent for effecting the primer extension reaction (called herein "agent for polymerization"), and the reaction is allowed to occur under conditions known in the art. The agent for polymerization may also be added together with the other reagents if it is heat stable. This synthesis (or amplification) reaction may occur at room temperature up to a temperature above which the agent for polymerization no longer functions. Thus, for example, if DNA polymerase is used as the agent, the temperature is generally no greater than about 40°C. Most conveniently the reaction occurs at room temperature. When using thermostable DNA polymerase such as Taq, higher temperature may be used.

The allele specific oligonucleotide primers are useful in determining whether a subject is at risk of having breast or ovarian cancer, and also useful for characterizing a tumor. Primers direct amplification of a target polynucleotide prior to sequencing. These unique BRCA2 oligonucleotide primers for exons 2-27 shown in TABLE II were designed and produced specifically to optimize amplification of portions of BRCA2 which are to be sequenced.

The primers used to carry out this invention embrace oligonucleotides of sufficient length and appropriate sequence to provide initiation of polymerization. Environmental conditions conducive to synthesis include the presence of nucleoside triphosphates and an agent for polymerization, such as DNA polymerase, and a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency in amplification, but may be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent for polymerization. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The oligonucleotide primer typically contains 18-28 bp plus in some cases an M13 "tail" for convenience.

Primers used to carry out this invention are designed to be substantially complementary to each strand of the genomic locus to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands under conditions which allow the agent for polymerization to perform. In other words, the primers should have sufficient complementarity with the 5' and 3' sequences flanking the mutation to hybridize therewith and permit amplification of the genomic locus.

Oligonucleotide primers of the invention are employed in the amplification process which is an enzymatic chain reaction that produces exponential quantities of polymorphic locus relative to the number of reaction steps involved. Typically, one primer is complementary to the negative (-) strand of the polymorphic locus and the other is complementary to the positive (+) strand. Accealing the primers to denatured nucleic acid followed by extension with an enzyme, such as the large fragment of DNA polymerase I (Klenow) and nucleotides, results in newly synthesized + and strands containing the target polymorphic locus sequence. Because these newly synthesized sequences are also templates, repeated cycles of denaturing, primer annealing, and extension results in exponential production of the region (i.e., the target polymorphic locus sequence) defined by the primers. The product of the chain reaction is a discreet nucleic acid

duplex with termini corresponding to the ends of the specific primers employed.

5

10

15

20

25

30

The oligonucleotide primers of the invention may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage, et al., Tetrahedron Letters, 22:1859-1862, 1981. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066.

The agent for polymerization may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase, polymerase muteins, reverse transcriptase, other enzymes, including heat-stable enzymes (*i.e.*, those enzymes which perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation), such as *Taq* polymerase. Suitable enzymes will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each polymorphic locus nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths.

The newly synthesized strand and its complementary nucleic acid strand will form a double-stranded molecule under hybridizing conditions described above and this hybrid is used in subsequent steps of the process. In the next step, the newly synthesized double-stranded molecule is subjected to denaturing conditions using any of the procedures described above to provide single-stranded molecules.

The steps of denaturing, annealing, and extension product synthesis can be repeated as often as needed to amplify the target polymorphic locus nucleic acid sequence to the extent necessary for detection. The amount of the specific nucleic acid sequence produced will accumulate in an

5

10

15

20

25

30

exponential fashion. Amplification is described in *PCR. A Practical Approach*, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, 1992.

The amplification products may be detected by Southern blots analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing a very low level of the nucleic acid sequence of the polymorphic locus is amplified, and analyzed via a Southern blotting technique or similarly, using dot blot analysis. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal. Alternatively, probes used to detect the amplified products can be directly or indirectly detectably labeled, for example, with a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme. Those of ordinary skill in the art will know of other suitable labels for binding to the probe, or will be able to ascertain such, using routine experimentation.

Sequences amplified by the methods of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific DNA sequence such as PCR, oligomer restriction (Saiki, et.al., Bio/Technology, 3:1008-1012, 1985), allele-specific oligonucleotide (ASO) probe analysis (Conner, et al., Proc. Natl. Acad. Sci. U.S.A., 80:278, 1983), oligonucleotide ligation assays (OLAs) (Landgren, et al., Science, 241:1007, 1988), and the like. Molecular techniques for DNA analysis have been reviewed (Landgren, et al., Science, 242:229-237, 1988).

Preferably, the method of amplifying is by PCR, as described herein and as is commonly used by those of ordinary skill in the art. Alternative methods of amplification have been described and can also be employed as long as the BRCA2 locus amplified by PCR using primers of the invention is similarly amplified by the alternative means. Such alternative amplification systems include but are not limited to self-sustained sequence replication, which begins with a short sequence of RNA of interest and a T7 promoter. Reverse transcriptase copies the RNA into cDNA and degrades the RNA,

5

10

15

20

25

30

followed by reverse transcriptase polymerizing a second strand of DNA. Another nucleic acid amplification technique is nucleic acid sequence-based amplification (NASBA) which uses reverse transcription and T7 RNA polymerase and incorporates two primers to target its cycling scheme. NASBA can begin with either DNA or RNA and finish with either, and amplifies to 10⁸ copies within 60 to 90 minutes. Alternatively, nucleic acid can be amplified by ligation activated transcription (LAT). LAT works from a single-stranded template with a single primer that is partially single-stranded and partially double-stranded. Amplification is initiated by ligating a cDNA to the promoter oligonucleotide and within a few hours, and amplification is 108 to 10⁹ fold. Another amplification system useful in the method of the invention is the Qβ Replicase System. The Qβ replicase system can be utilized by attaching an RNA sequence called MDV-1 to RNA complementary to a DNA sequence of interest. Upon mixing with a sample, the hybrid RNA finds its complement among the specimen's mRNAs and binds, activating the replicase to copy the tag-along sequence of interest. Another nucleic acid amplification technique, ligase chain reaction (LCR), works by using two differently labeled halves of a sequence of interest which are covalently bonded by ligase in the presence of the contiguous sequence in a sample, forming a new target. The repair chain reaction (RCR) nucleic acid amplification technique uses two complementary and target-specific oligonucleotide probe pairs, thermostable polymerase and ligase, and DNA nucleotides to geometrically amplify targeted sequences. A 2-base gap separates the oligonucleotide probe pairs, and the RCR fills and joins the gap, mimicking normal DNA repair. Nucleic acid amplification by strand displacement activation (SDA) utilizes a short primer containing a recognition site for hincll with short overhang on the 5' end which binds to target DNA. A DNA polymerase fills in the part of the primer opposite the overhang with sulfur-containing adenine analogs. Hincll is added but only cuts the unmodified DNA strand. A DNA polymerase that lacks 5' exonuclease activity enters at the site of the nick and begins to polymerize, displacing the initial primer strand downstream and building a new one which serves as

5

10

15

20

25

30

more primer. SDA produces greater than 10⁷-fold amplification in 2 hours at 37°C. Unlike PCR and LCR, SDA does not require instrumented Temperature cycling.

Another method is a process for amplifying nucleic acid sequences from a DNA or RNA template which may be purified or may exist in a mixture of nucleic acids. The resulting nucleic acid sequences may be exact copies of the template, or may be modified. The process has advantages over PCR in that it increases the fidelity of copying a specific nucleic acid sequence, and it allows one to more efficiently detect a particular point mutation in a single assay. A target nucleic acid is amplified enzymatically while avoiding strand displacement. Three primers are used. A first primer is complementary to the first end of the target. A second primer is complementary to the second end of the target. A third primer which is similar to the first end of the target and which is substantially complementary to at least a portion of the first primer such that when the third primer is hybridized to the first primer, the position of the third primer complementary to the base at the 5' end of the first primer contains a modification which substantially avoids strand displacement. This method is detailed in U.S. Patent 5,593,840 to Bhatnagar et al. 1997, incorporated herein by reference.

Finally, recent application of DNA chips or microarray technology where DNA or oligonucleotides are immobilized on small solid support may also be used to rapidly sequence sample BRCA2 gene and analyze its expression. Typically, high density arrays of DNA fragment are fabricated on glass or nylon substrates by *in situ* light-directed combinatorial synthesis or by conventional synthesis followed by immobilization (Fodor *et al.* U.S. patent No. 5,445,934). Sample DNA or RNA may be amplified by PCR, labeled with a fluorescent tag, and hybridized to the microarray. Examples of this technology are provided in U.S. Patents 5,510, 270, U.S. 5,547,839, incorporated herein by reference.

All exonic and adjacent intronic sequences of the BRCA2 gene were obtained by end to end sequencing of five normal subjects in the manner described above followed by analysis of the data obtained. The data obtained provided us with the opportunity to establish the correct

intronic/exonic structure of the BRCA2 gene. In addition, we evaluated six previously published normal polymorphisms (1342, 2457, 3199, 3624, 4035, and 7470) for correctness and frequency in the population, and to identify four additional polymorphisms not previously characterized (1093, 1593, 2908, and 9079).

GENE THERAPY

5

10

15

20

25

30

The polynucleotide(s) which result from either sense or antisense transcription of any exon or the entire coding sequence or fragments of BRCA2 gene may be used for gene therapy. A variety of methods are known for gene transfer, any of which might be available for use.

Direct injection of Recombinant DNA in vivo:

- 1. Direct injection of "naked" DNA directly with a syringe and needle into a specific tissue, infused through a vascular bed, or transferred through a catheter into endothelial cells.
- 2. Direct injection of DNA that is contained in artificially generated lipid vesicles or other encapsulating vehicles.
- 3. Direct injection of DNA conjugated to a target receptor structure, such as a diptheria toxin, an antibody or other suitable receptor.
- 4. Direct injection by particle bombardment. For example, the DNA may be coated onto gold particles and shot into the cells.

Human Artificial Chromosomes

The gene delivery approach involves the use of human chromosomes that have been stripped down to contain only the essential components for replication and the genes desired for transfer.

keceptor-Mediated Gene Transfer

DNA is linked to a targeting molecule that will bind to specific cell-surface receptors, inducing endocytosis and transfer of the DNA into mammalian cells. One such technique uses poly-L-lysine to link asialoglycoprotein to DNA. An adenovirus is also added to the complex to disrupt the lysosomes and thus allow the DNA to avoid degradation and move to the nucleus. Infusion of these particles intravenously has resulted in gene transfer into hepatocytes.

RECOMBINANT VIRUS VECTORS

Several vectors may be used in gene therapy. Among them are the Moloney Murine Leukemia Virus (MoMLV) Vectors, the adenovirus vectors, the Adeno-Associated Virus (AAV) vectors, the herpes simplex virus (HSV) vectors, the poxvirus vectors, the retrovirus vectors, and human immunodeficiency virus (HIV) vectors.

GENE REPLACEMENT AND REPAIR

The ideal genetic manipulation for treatment of a genetic disease would be the actual replacement of the defective gene with a normal copy of the gene. Homologous recombination is the term used for switching out a section of DNA and replacing it with a new piece. By this technique, the defective gene may be replaced with a normal gene which expresses a functioning BRCA2 tumor growth inhibitor protein.

A complete description of gene therapy can also be found in "Gene Therapy A Primer For Physicians" 2d Ed. by Kenneth W. Culver, M.D. Publ. Mary Ann Liebert Inc. (1996). Two Gene Therapy Protocols for BRCA1 gene have been approved by the Recombinant DNA Advisory Committee for Jeffrey T. Holt et al. They are listed as 9602-148, and 9603-149 and are available from the NIH. Protocols for BRCA2 gene therapy may be similarly employed. The isolated BRCA2 gene may be synthesized or constructed from amplification products and inserted into a vector such as the LXSN vector.

25

30

5

10

15

20

A BRCA2 POLYPEPTIDE OR ITS FUNCTIONAL EQUIVALENT

The growth of breast and ovarian cancer may be arrested or prevented by directly increasing the BRCA2 protein level where inadequate functional BRCA2 activity is responsible for breast and ovarian cancer. The cDNA and amino acid sequences of five novel BRCA2 haplotypes are disclosed herein (SEQ ID No:4-13). All or a fragment of BRCA2 protein may be used in therapeutic or prophylactic treatment of breast and ovarian cancer. Such a fragment may have a similar biological function as the native BRCA2 protein

5

10

15

20

25

30

or may have a desired biological function as specified below. BRCA2 polypeptides or their functional equivalents including homologous and modified polypeptide sequences are also within the scope of the present invention. Changes in the native sequence may be advantageous in producing or using the BRCA2 derived polypeptides or functional equivalents suitable for therapeutic or prophylactic treatment of breast and ovarian cancer. For example, these changes may be desirable for producing resistance against *in vivo* proteolytic cleavage, for facilitating transportation and delivery of therapeutic reagents, for localizing and compartmentalizing tumor suppressing agents, or for expression, isolating and purifying the target species.

There are a variety of methods to produce an active BRCA2 polypeptide or a functional equivalent as a tumor growth inhibitor. For example, one or more amino acids may be substituted, deleted, or inserted using methods well known in the art (Maniatis et al., 1982). Considerations of polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphiphathic nature of the amino acids play an important role in designing homologous polypeptide changes suitable for the intended treatment. In particular, conservative amino acid substitution using amino acids that are related in side-chain structure and charge may be employed to preserve the chemical and biological property. A homologous polyeptide typically contains at least 70% homology to the native sequence. Unnatural forms of the polypeptide may also be incorporated so long as the modification retains substantial biological activity. These unnatural polypeptides typically include structural mimics and chemical medications, which have similar threedimensional structures as the active regions of the native BRCA2 protein. For example, these modifications may include terminal D-amino acids, cyclic peptides, unnatural amino acids side chains, pseudopeptide bonds, Nterminal acetylation, glycosylation, and biotinylation, etc. These unnatural forms of polypeptide may have a desired biological function, for example, they may be particularly robust in the presence of cellular or serum proteases and exopeptidase. An effective BRCA2 polypeptide or a functional equivalent may also be recognized by the reduction of the native

5

10

15

20

25

30

BRCA2 protein. Regions of the BRCA2 protein may be systematically deleted to identify which regions are essential for tumor growth inhibitor activity. These smaller fragments of BRCA2 protein may then be subjected to structural and functional modification to derive therapeutically or prophylactically effective regiments. Finally, drugs, natural products or small molecules may be screened or synthesized to mimic the function of the BRCA2 protein. Typically, the active species retain the essential threedimensional shape and chemical reactivity, and therefore retain the desired aspects of the biological activity of the native BRCA2 protein. The activity and function of BRCA2 may include transactivation, granin, DNA repair among others. Functions of BRCA2 protein are also reviewed in Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998) and Zhang et al., Cell 92:433-436 (1998). It will be apparent to one skilled in the art that a BRCA2 polypeptide or a functional equivalent may be selected because such polypeptide or functional equivalent possesses similar biological activity as the native BRCA2 protein.

EXPRESSION OF THE BRCA2 PROTEIN AND POLYPEPTIDE IN HOST CELLS

All or fragments of the BRCA2 protein and polypeptide may be produced by host cells that are capable of directing the replication and the expression of foreign genes. Suitable host cells include prokaryotes, yeast cells, or higher eukaryotic cells, which contain an expression vector comprising all or a fragment of the BRCA2 cDNA sequence (SEQ. ID No: 4, 6, 8, 10, or 12) operatively linked to one or more regulatory sequences to produce the intended BRCA2 protein or polypeptide. Prokaryotes may include gram negative or gram positive organisms, for example *E. coli* or *Bacillus* strains. Suitable eukaryotic host cells may include yeast, virus, and mamalian systems. For example, Sf9 insect cells and human cell lines, such as COS, MCF7, HeLa, 293T, HBL100, SW480, and HCT116 cells.

A broad variety of suitable expression vectors are available in the art.

An expression vector typically contains an origin of replication, a promoter, a phenotypic selection gene (antibiotic resistance or autotrophic requirement),

5

10

15

20

25

30

and a DNA sequence coding for all or fragments of the BRCA2 protein. The expression vectors may also include other operatively linked regulatory DNA sequences known in the art, for example, stability leader sequences, secretory leader sequences, restriction enzyme cleavage sequences, polyadenylation sequences, and termination sequences, among others. The essential and regulatory elements of the expression vector must be compatible with the intended host cell. Suitable expression vectors containing the desired coding and control regions may be constructed using standard recombinant DNA techniques known in the art, many of which are described in Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). For example, suitable origins of replication may include Col E1, SV4O viral and M13 origins of replication. Suitable promoters may be constitutive or inducible, for example, tac promoter, lac Z promoter, SV40 promoter, MMTV promoter, and LXSN promoter. Examples of selectable markers include neomycin, ampicillin, and hygromycin resistance and the like. Many suitable prokaryotic, viral and mammalian expression vectors may be obtained commercially, for example, from Invitrogen Corp., San Diego, CA or from Clontech, Palo Alto, CA. It may be desirable that the BRCA2 protein or polypeptide is produced as a fusion protein to enhance the expression in selected host cells, to detect the expression in transfected cells, or to simplify the purification process. Suitable fusion partners for the BRCA2 protein or polypeptide are well known in the art and may include β -galactosidase, glutathione-S-transferase, and poly-histidine tag.

Expression vectors may be introduced into host cells by various methods known in the art. The transformation procedure used depends upon the host to be transformed. Methods for introduction of vectors into host cells may include calcium phosphate precipitation, electrosporation, dextranmediated transfection, liposome encapsulation, nucleus microinjection, and viral or phage infection, among others.

Once an expression vector has been introduced into a suitable host cell, the host cell may be cultured under conditions permitting expression of large amounts of the BRCA2 protein or polypeptide. The expression product

5

10

15

20

25

30

may be identified by many approaches well known in the art, for example, sequencing after PCR-based amplification, hybridization using probes complementary to the desired DNA sequence, the presence or absence of marker gene functions such as enzyme activity or antibiotic resistance, the level of mRNA production encoding the intended sequence, immunological detection of a gene product using monoclonal and polyclonal antibodies, such as Western blotting or ELISA. The BRCA2 protein or polypeptides produced in this manner may then be isolated following cell lysis and purified using various protein purification techniques known in the art, for example, ion exchange chromatography, gel filtration chamber and immunoaffinity chromatography.

It is generally preferred that whenever possible, longer fragments of BRCA2 protein or polypeptide are used, particularly to include the desired functional domains of BRCA2 protein. Expression of shorter fragments of DNA may be useful in generating BRCA2 derived immunogen for the production of anti-BRCA2 antibodies. It should, of course, be understood that not all expression vectors, DNA regulatory sequences or host cells will function equally well to express the BRCA2 protein or polypeptides of the present invention. However, one of ordinary skill in the art may make a selection among expression vectors, DNA regulatory sequences, host cells, and codon usage in order to optimize expression using known technology in the art without undue experimentation. Studies of BRCA2 protein function and examples of genetic manipulation of BRCA2 protein are summarized in two recent review articles, Bertwistle and Ashworth, *Curr. Opin. Genet. Dev.* 8(1): 14-20 (1998) and Zhang *et al.*, *Cell* 92:433-436 (1998).

IN VITRO SYNTHESIS AND CHEMICAL SYNTHESIS

Although it is preferred that fragments of the BRCA2 protein or polypeptides be obtained by overexpression in prokaryotic or eukaryotic host cells, the BRCA2 polypeptides or their functional equivalents may also be obtained by *in vitro* translation or synthetic means by methods known to those of ordinary skill in the art. For example, *in vitro* translation may employ an mRNA encoded by a DNA sequence coding for fragments of the BRCA2

protein or polypeptides. Chemical synthesis methodology such as solid phase synthesis may be used to synthesize a BRCA2 polypeptide structural mimic and chemically modified analogs thereof. The polypeptides or the modifications and mimic thereof produced in this manner may then be isolated and purified using various purification techniques, such as chromatographic procedures including ion exchange chromatography, gel filtration chromatography and immunoaffinity chromatography.

PROTEIN REPLACEMENT THERAPY

10

15

20

25

30

The tumor suppressing function of BRCA2 suggests that various BRCA2 protein targeted therapies may be utilized in treating and preventing tumors in breast and ovarian cancer. The present invention therefore includes therapeutic and prophylactic treatment of breast and ovarian cancer using therapeutic pharmaceutical compositions containing the BRCA2 protein, polypeptides, or their functional equivalents. For example, protein replacement therapy may involve directly administering the BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent in a pharmaceutically effective carrier. Alternatively, protein replacement therapy may utilize tumor antigen specific antibody fused to fragments of the BRCA2 protein, a polypeptide, or a functional equivalent to deliver anti-cancer regiments specifically to the tumor cells.

To prepare the pharmaceutical compositions of the present invention, an active BRCA2 protein, a BRCA2 polypeptide, or its functional equivalent is combined with a pharmaceutical carrier selected and prepared according to conventional pharmaceutical compounding techniques. A suitable amount of the composition may be administered locally to the site of a tumor or systemically to arrest the proliferation of tumor cells. The methods for administration, may include parenteral, oral, or intravenous, among others according to established protocols in the art.

Pharmaceutically acceptable solid or liquid carriers or components which may be added to enhance or stabilize the composition, or to facilitate preparation of the composition include, without limitation, syrup, water, isotonic solution, 5 % glucose in water or buffered sodium or ammonium

acetate solution, oils, glycerin, alcohols, flavoring agents, preservatives, coloring agents, starches, sugars, diluents, granulating agents, lubricants, binders, and sustained release materials. The dosage at which the therapeutic compositions are administered may vary within a wide range and depends on various factors, such as the stage of cancer progression, the age and condition of the patient, and may be individually adjusted.

DIAGNOSTIC REAGENTS

5

10

15

20

25

30

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may be used in a wide variety of ways in addition to gene therapy and protein replacement therapy. They may be useful as diagnostic reagents to measure normal or abnormal activity of BRCA2 at the DNA, RNA, and protein level. The present invention therefore encompasses the diagnostic reagents derived from the BRCA2 cDNA and protein sequences as set forth in SEQ. ID. Nos: 4-13. These reagents may be utilized in methods for monitoring disease progression, for determining patients suited for gene and protein replacement therapy, or for detecting the presence or quantifying the amount of a tumor growth inhibitor following such therapy. Such methods may involve conventional histochemical techniques, such as obtaining a tumor tissue from the patient, preparing an extract and testing this extract for tumor growth or metabolism. For example, the test for tumor growth may involve measuring abnormal BRCA2 activity using conventional diagnostic assays, such as Southern, Northern, and Western blotting, PCR, RT-PCR, and immunoprecipitation. In biopsies of tumor tissues, the loss of BRCA2 expression in tumor tissue may be verified by RT-PCR and Northern blotting at the RNA level. A Southern blot analysis, genomic PCR, or fluorescence in situ hybridization (FISH) may also be performed to examine the mutations of BRCA2 at the DNA level. And, a Western blotting, protein truncation assay, or immunoprecipitation may be utilized to analysis the effect at the protein level.

These diagnostic reagents are typically either covalently or non convalently attached to a detectable label. Such a label includes a radioactive label, a colorimetric enzyme label, a fluorescence label, or an

epitope label. Frequently, a reporter gene downstream of the regulatory sequences is fused with the BRCA2 protein or polypeptide to facilitate the detection and purification of the target species. Commonly used reporter genes in BRCA2 fusion proteins include β -galactosidase and luciferase gene.

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 in oncogenesis or subcellular localization of BRCA2 protein in normal and cancerous cell. For example, yeast two-hybrid system has been used in the study of cellular function of BRCA2 to identify the regulator and effector of BRCA2 tumor suppressing function (Sharan et al., Nature 386:804-810 (1997) and Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988)). In addition, the BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be used in *in vivo* cell based and *in vitro* cell free assays to screen natural products and synthetic compounds which may mimic, regulate or stimulate BRCA2 protein function.

ANTISENSE INHIBITION

5

10

15

20

25

30

Antisense suppression of endogenous BRCA2 expression may assess the effect of BRCA2 protein on cell growth inhibition using known method in the art (Crooke, *Annu. Rev. Pharmacol. Toxicol.* 32:329-376 (1992) and Robinson-Benion and Holt, *Methods Enzymol.* 254:363-375 (1995)). Given the cDNA sequence as set forth in SEQ ID. NO: 4, 6, 8, 10, and 12, one of skill in the art can readily obtain anti-sense strand of DNA and RNA sequences to interfere with the production of wild-type BRCA2 protein or the mutated form of BRCA2 protein. Alternatively, antisense oligonucleotide may be designed to target the control sequences of BRCA2 gene to reduce or prevent the expression of the endogenous BRCA2 gene.

ANTIBODIES

The BRCA2 protein, polypeptides, or their functional equivalents may be used as immunogens to prepare polyclonal or monoclonal antibodies

capable of binding the BRCA2 derived antigens in a known manner (Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988). These antibodies may be used for the detection of the BRCA2 protein, polypeptides, or a functional equivalent in an immunoassay, such as ELISA, Western blot, radioimmunoassay, enzyme immunoassay, and immunocytochemistry. Typically, an anti-BRCA2 antibody is in solution or is attached to a solid surface such as a plate, a particle, a bead, or a tube. The antibody is allowed to contact a biological sample or a blot suspected of containing the BRCA2 protein or polypeptide to form a primary immunocomplex. After sufficient incubation period, the primary immunocomplex is washed to remove any non-specifically bound species. The amount of specifically bound BRCA2 protein or polypeptide may be determined using the detection of an attached label or a marker, such as a radioactive, a fluorescent, or an enzymatic label. Alternatively, the detection of BRCA2 derived antigen is allowed by forming a secondary immunocomplex using a second antibody which is attached with a such label or marker. The antibodies may also be used in affinity chromatography for isolating or purifying the BRCA2 protein, polypeptides or their functional equivalents.

20

25

30

15

5

10

EXAMPLE 1

<u>Determination of the Coding Sequence Haplotypes of the BRCA2 Gene</u> <u>From Normal Individuals</u>

Approximately 150 volunteers were screened in order to identify individuals with no cancer history in their immediate family (i.e. first and second degree relatives). Each person was asked to fill out a hereditary cancer prescreening questionnaire (See TABLE I). Five of these were randomly chosen for end-to-end sequencing of their BRCA2 gene. A first degree relative is a parent, sibling, or offspring. A second degree relative is an aunt, uncle, grandparent, grandchild, niece, nephew, or half-sibling.

Genomic DNA was isolated from white blood cells of five normal subjects selected from analysis of their answers to the questions above.

Dideoxy sequence analysis was performed following polymerase chain reaction amplification.

All exons of the BRCA2 gene were subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye was attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer® cat# 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator" purchased through ABI.

1. Polymerase Chain Reaction (PCR) Amplification

5

10

20

25

30

Genomic DNA (100 nanograms) extracted from white blood cells of five normal subjects. Each of the five samples was sequenced end to end. Each sample was amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl₂), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer, 2.5 microliters reverse primer, and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The primers in TABLE II below were used to carry out amplification of the various sections of the BRCA2 gene samples. The primers were synthesized on an DNA/RNA Synthesizer Model 394[®].

Thirty-five cycles were performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time was increased to 5 minutes, and during the last cycle in which the extension time was increased to 5 minutes.

PCR products were purified using Qia-quick[®] PCR purification kits (Qiagen[®], cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD₂₆₀ on a Beckman DU 650 spectrophotometer.

5

10

15

20

25

30

2. Dideoxy Sequence Analysis

Fluorescent dye was attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer[®] cat # 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator[®]" purchased through ABI.

3. RESULTS

Based upon the sequencing of the five normal individuals, it was determined that the standard sequence found in both GenBank and BIC were inaccurate. In Genbank, a 10 bp stretch (5'-TTTATTTTAG-3') was mistakenly listed as exonic at the 5' end of exon 5 while it should be intronic which would not be included in the cDNA and resultant protein. In addition, a more detrimental error that has the significant potential to lead to an incorrect diagnosis of breast cancer propensity exists in both Genbank and BIC: a sequence of 16 bp (5'-GTGTTCTCATAAACAG-3') should be at the end of exon 15, but instead is listed at the beginning of exon 16 in the database. The disclosure and listing of GenBank is shown in Figure 1. The correct intron/exon sequence of BRCA2 is presented in Figure 2, wherein,

- (1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5 (corrected exon 5 is listed as SEQ. ID. NO: 1) and
- (2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon 16 (corrected exons 15 and 16 are listed as SEQ. ID. No: 2 and 3 respectively)

The BIC BRCA2 sequence also contains sequence errors in which a strech of nine nucleotides at positions 5554-5460 is listed as CGTTTGTGT (amino acids: Arg-Leu-Cys). The correct sequence at these positions is GTTTGTGTT (amino acids: Val-Cys-Val). In addition, the BIC BRCA2 nuclotides at positions 2024 (codon 599), 4553 (codon 1442), 4815 (codon 1529), 5841 (codon 1871), and 5972 (codon 1915) are T, T, A, C, and T respectively, wherein the correct nucleotides at these positions are C, C, G, T, and C respectively. Among them, the nuclotide errors at codon 599, 1442, 1915 result in amino acids changes.

5

10

15

20

25

Additional differences in the nucleic acids of the five normal individuals were found in ten polymorphic locations. The changes and their positions are found in TABLE III. The individual haplotypes of each chromosome of BRCA2 are displayed in FIGURE 3. In each case, the initial haplotype reported in Genbank (accession number U43746) was subtracted to determine the new haplotypes OMI 1-5. Thus, the Genbank sequence only represents 50% of the haplotypes found; the five new BRCA2 (ormi 1-5) DNA sequences are shown as SEQ. ID. NO: 4, 6, 8, 10, and 12, respectively (See FIGURE 3), and the corresponding polypeptides are listed as SEQ. ID. NO: 5, 7, 9, 11, and 13 respectively. In combination, these seven haplotypes represent a functional allele profile for the BRCA2 gene.

The data show that for each of the samples, all exons of BRCA2 were identical except in the region of ten polymorphisms. Six of these polymorphisms were previously identified (Tartigan *et al.*, *Nature Genetics* 12: 333-337 (1996); Phelan *et al.*, *Nature Genetics* 13: 120-122 (1996); Couch *et al.*, *Nature Genetics* 13: 123-125 (1996); Teng, *et al.*, *Nature Genetics* 13: 241-244 (1996); Schubert *et al* 60: 1031-1040 (1997)), but four were unique to this work. Even though the individual polymorphisms may have been identified, none of these complete haplotypes has been previously determined.

TABLE I

Hereditary Cancer Pre-Screening Questionnaire

Part A: Answer the following questions about your family

5

30

40

45

50

1. To your knowledge, has anyone in your family been diagnosed with a very specific hereditary colon disease called Familial Adenomatous Polyposis (FAP)?

- 2. To your knowledge, have you or any aunt had breast cancer diagnosed before the age 35?
 - 3. Have you had Inflammatory Bowel Disease, also called Crohn's Disease or Ulcerative Colitis, for more than 7 years?
- 15 Part B: Refer to the list of cancers below for your responses only to questions in Part B

	Bladder Cancer	Lung Cancer	Pancreatic Cancer
	Breast Cancer	Gastric Cancer	Prostate Cancer
	Colon Cancer	Malignant Melanoma	Renal Cancer
20	Endometrial Cancer	Ovarian Cancer	Thyroid Cancer

- 4. Have your mother or father, your sisters or brothers or your children had any of the listed cancers?
- Have there been diagnosed in your <u>mother</u>'s brothers or sisters, or your <u>mother</u>'s parents <u>more than one</u> of the cancers in the above list?
 - 6. Have there been diagnosed in your <u>father</u>'s brothers or sisters, or your <u>father</u>'s parents <u>more than one</u> of the cancers in the above list?

Part C: Refer to the list of relatives below for responses only to questions in Part C

You Your mother
Your sisters or brothers

S & uncles)
Your children
Your mother's sisters or brothers (maternal aunts)
Your mother's parents (maternal grandparents)

- Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer?
 Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 8. Is there a total of 4 or more of any cancers in the list of relatives above other than "simple" skin cancers?

Part D: Refer to the list of relatives below for responses only to questions in Part U.

You Your father
Your sisters or brothers Your father's sisters or brothers (paternal aunts and uncles)
Your children Your father's parents (paternal grandparents)

- 9. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer?

 Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 10. Is there a <u>total of 4 or more</u> of any cancers in the list of relatives above other than "simple" skin cancers?
 - © Copyright 1996, OncorMed, Inc.

NOT FURNISHED UPON FILING

NO PRESENTADO/(A) EN EL MOMENTO DE LA PRESENTACIÓN

NON SOUMIS(E) AU MOMENT DU DEPÔT

NOT FURNISHED UPON FILING

NO PRESENTADO/(A) EN EL MOMENTO DE LA PRESENTACIÓN

NON SOUMIS(E) AU MOMENT DU DEPÔT

TABLE II BRCA2 PRIMER SEQUENCES

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Oligo Length	PCR Product Length	SEQ. ID. Number
2	BRCA2-2F	5'-TGA GTT TTA CCT CAG TCA CA-3'	20	263	14
7	BRCA2-2R/M 13R	5'-CAG GAA ACA GCT ATG ACC CTG TGA CGT ACT GGG TTT TTA GC-3'	41		15
ဇ	BRCA2-3FII	3'-GAT CTT TAA CTG TTC TGG GTC ACA-3'	24	364	16
က	BRCA2-3RII	5'-CCC AGC ATG ACA CAA TTA ATG A-3'	22		17
4	BRCA2-4F/M 13F	5'-TGT AAA ACG ACG GCC AGT AGA ATG CAA ATT TAT AAT CCA GAG TA-3'	44	268	18
4	BRCA2-4R-1A	5'-ATC AGA TTC ATC TTT ATA GAA C-3'	22		19
5&6	BRCA2-5+6F/M13F	5-TGT AAA ACG ACG GCC AGT TGT GTT GGC ATT TTA AAC ATC A-3'	40	453	20
586	BRCA2-5+6R/M13R	5-CAG GAA ACA GCT ATG ACC CAG GGC AAA GGT ATA ACG CT-3'	38		21
7	BRCA2-7F/M13F	5'-TGT AAA ACG ACG GCC AGT TAA GTG AAA TAA AGA GTG AA-3'	38	248	22
7	BRCA2-7R/M13R	5'-CAG GAA ACA GCT ATG ACC AGA AGT ATT AGA GAT GAC-3'	36		23
80	BRCA2-8F/M13F	5'-TGT AAA ACG ACG ACT GCC ATA TCT TAC CAC CTT GTG A-3'	40	319	24
	BRCA2-8FIA	5-TTG CAT TCT AGT GAT AAT ATA C-3'	22	143	25
8	BRCA2-8RIA	5'-AAT TGT TAG CAA TTT CAA C-3'	19		56
6	BRCA2-9F/M13F	5-TGT AAA ACG ACG GCC AGT TGG ACC TAG GTT GAT TGC AGA T-3'	40	338	27
6	BRCA2-9R/M13R	5-CAG GAA ACA GCT ATG ACC TAA ACT GAG ATC ACG GGT GAC A-3'	40		28
10A	BRCA2-10AF	5-GAA TAA TAT AAA TTA TAT GGC TTA-3'	24	255	59
10A	BRCA2-10AR/M13R	5-CAG GAA ACA GCT ATG ACC CCT AGT CTT GCT AGT TCT T-3'	37		30
10B	BRCA2-10BF/M13F	5- TGT AAA ACG ACG GCC AGT ARC TGA AGT GGA ACC AAA TGA TAC-3'	42	621	31

TABLE II BRCA2 PRIMER SEQUENCES

Fxon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED	Oligo	PCR	SEQ. ID.
		M13 FORWAR M13 REVERS	Length	Product Length	Number
10B	BRCA2-10BR/M13R	5- CAG GAA ACA GCT ATG ACC ACG TGG CAA AGA ATT CTC TGA AGT AA-3'	44		32
10C	BRCA2-10CF/M13F	5-TGT AAA ACG ACG GCC AGT CAG CAT CTT GAA TCT CAT ACA G-3'	40	508	33
10C	BRCA2-10CRII	5-AGA CAG AGG TAC CTG AAT C-3'	19		8
11	BRCA2-11AF-M13	5- TGT AAA ACG ACG GCC AGT TGG TAC TTT AAT TIT GTC ACT T-3'	40	304	35
11	BRCA2-11AR-M13	5-CAG GAA ACA GCT ATG ACC TGC AGG CAT GAC AGA GAA T-3'	37		36
11	BRCA2-11BF	5'-AAG AAG CAA AAT GTA ATA AGG A-3'	22	411	37
-	BRCA2-11BR	5'-CAT TTA AAG CAC ATA CAT CTT G-3'	22		38
11	BRCA2-11CF	5'-TCT AGA GGC AAA GAA TCA TAC-3'	21	349	39
11	BRCA2-11CR	5-CAA GAT TAT TCC TTT CAT TAG C-3'	22		40
11	BRCA2-11DF	5-AAC CAA AAC ACA AAT CTA AGA G-3'	22	344	41
1-1	BRCA2-11DR	5-GTC ATT TIT ATA TGC TGC TIT AC-3'	23		42
17	BRCA2-11EF	5'-GGT TTT ATA TGG AGA CAC AGG-3'	21	369	43
+	BRCA2-11ER	5-GTA TTT ACA ATT TCA ACA CAA GC-3'	23		44
11	BRCA2-11FF	5'-ATC ACA GTT TTG GAG GTA GC-3'	20	368	45
11	BRCA2-11FR	5'-CTG ACT TCC TGA TTC TTC TAA-3'	21		46
11	BRCA2-11GF	5'-CTC AGA TGT TAT TTT CCA AGC-3'	21	366	47
11	BRCA2-11GR	5-CTG TTA AAT AAC CAG AAG CAC-3'	21		48
11	BRCA2-11HF	5-AGG TAG ACA GCA AGC-3'	18	360	49
11	BRCA2-11HR	5'-GTA ATA TCA GTT GGC ATT TAT T-3'	22		20
11	BRCA2-11IF	5-TGC AGA GGT ACA TCC AAT AAG-3'	21	326	51

TABLE II BRCA2 PRIMER SEQUENCES

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG ACC M13 REVERSE = CAG GAA ACA GCT ATG ACC	Oligo	PCR Product Length	SEQ. ID. Number
11	BRCA2-11IR	5'-GAT CAG TAA ATA GCA AGT CCG-3'	21		52
11	BRCA2-11JF	5'-TAC TGA AAA TGA AGA TAA CAA AT-3'	23	477	53
17	BRCA2-11JR	5'-ATT TTG TTC TTT CTT ATG TCA G-3'	22		\$
41	BRCA2-11KF-M13	5'-TGT AAA ACG ACG GCC AGT CTA CTA AAA CGG AGC AA-3'	35	382	55
11	BRCA2-11KR-M13	5'-CAG GAA ACA GCT ATG ACC GTA TGA AAA CCC AAC AG-3'	35		56
=	BRCA2-11LF	5'-CAC AAA ATA CTG AAA GAA AGT G-3'	22	374	22
11	BRCA2-11LR	5'-GGC ACC ACA GTC TCA ATA G-3'	19		28
11	BRCA2-11MF	5-GCA AAG ACC CTA AAG TAC AG-3'	50	409	59
11	BRCA2-11MR	5-CAT CAA ATA TTC CTT CTC TAA G-3'	22		09
11	BRCA2-11NF-M13	5'-TGT AAA ACG ACG GCC AGT GAA AAT TCA GCC TTA GC-3'	35	306	61
11	BRCA2-11NR-M13	5'- CAG GAA ACA GCT ATG ACC ATC AGA ATG GTA GGA AT-3'	35		62
-	BRCA2-110F	5'-GTA CTA TAG CTG AAA ATG ACA A-3'	22	383	63
11	BRCA2-110R	5'-ACC ACT GGC TAT CCT AAA TG-3'	20		2
17	BRCA2-11PF	5'-TGA AGA TAT TTG CGT TGA GG-3'	20	355	65
=	BRCA2-11PR	5'-GTC AGC AAA AAC CTT ATG TG-3'	20		99
=	BRCA2-11QF	5'-ACG AAA ATT ATG GCA GGT TGT-3'	21	337	29
=	BRCA2-11QR	5'-CTT GTC TTG CGT TTT GTA ATG-3'	21		89
=	BRCA2-11RF	5-GCT TCA TAA GTC AGT CTC AT-3'	20	360	69
=	BRCA2-11RR	5:-TCA AAT TCC TCT AAC ACT CC-3'	20		20
=	BRCA2-11SF-M13	6'-TGT AAA ACG ACG GCC AGT TAC AGC AAG TGG AAA GC-3'	35	458	71

TABLE II BRCA2 PRIMER SEQUENCES

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 PEVERSE = CAG GAA ACA GCT ATG ACC	Oligo Length	PCR Product Length	SEQ. ID. Number
=	BRCA2-11SR-M13	5-CAG GAA ACA GCT ATG ACC AAG TTT CAG TTT TAC CAA T-3'	37		72
=	BRCA2-11TF	5'-GTT CTT CAG AAA ATA ATC ACT C-3'	22	344	73
=	BRCA2-11TR	5-TGT AAA AAG AGA ATG TGT GGC-3'	21		74
: -	BRCA2-11UF-M13	5'-TGT AAA ACG ACG GCC AGT ACT TTT TCT GAT GTT CCT GTG-3'	39	328	75
=	BRCA2-11UR-M13	5'-CAG GAA ACA GCT ATG ACC TAA AAA TAG TGA TTG GCA ACA-3'	39		92
12	BRCA2-12F/M13F	5'-TGT AAA ACG ACG GCC AGT AGT GGT GTT TTA AAG TGG TCA AAA-3'	42	391	7.7
12	BRCA2-12R/M13R	5'-CAG GAA ACA GCT ATG ACC GGA TCC ACC TGA GGT CAG AAT A-3'	40		78
13	BRCA2/13-2F	5:-TAA CAT TTA AGC ATC CGT TAC-3'	21	310	62
13	BRCA2/13-2R	'5'-AAA CGA GAC TTT TCT CAT ACT GTA TTA G-3'	78		80
14	BRCA2-14F	5'-ACC ATG TAG CAA ATG AGG GTC T-3'	22	391	81
14	BRCA2-14AR	5'-GCT TTT GTC TGT TTT CCT CCA A-3'	22		82
15	BRCA2-15-2F	5'-CCA GGG GTT GTG CTT TTT AAA-3'	21	284	83
15	BRCA2-15FUT/M13-R 5'-CAG	5'-CAG GAA ACA GCT ATG ACC ACT CTG TCA TAA AAG CCA TC-3'	38		84
16	BRCA2-16AF	5:-TTT GGT TTG TTA TAA TTG TTT TTA-3'	24	394	85
16	BRCA2-16AR	5-CCA ACT TTT TAG TTC GAG AG-3'	50		86
17	BRCA2-17F	5-TTC AGT ATC ATC CTA TGT G-3'	19	282	87
17	BRCA2-17AR	5'-AGA AAC CTT AAC CCA TAC TG-3'	50		88
18	BRCA2-18FUT/M13-	5'-TGT AAA ACG ACG GCC AGT GAA TTC TAG AGT CAC ACT TCC-3'	33	275	89
18	BRCA2-18R/M13R	5'-CAG GAA ACA GCT ATG ACC TTT AAC TGA ATC AAT GAC TG-3'	38		06
19	BRCA2-19F/M13F	5'-TGT AAA ACG ACG GCC AGT AAG TGA ATA TTT TTA AGG CAG TT-3'	41	355	91

TABLE II BRCA2 PRIMER SEQUENCES

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT	Oligo	PCR Product Length	SEQ. ID. Number
10	RRCA2-19FUT/M13-R	RECA2-19FUT/M13-R 5'-CAG GAA ACA GCT ATG ACC AAG AGA CCG AAA CTC CAT CTC-3'	39		92
2 6	BRCA2-20F/M13F	5-TGT AAA ACG ACG GCC AGT CAC TGT GCC TGG CCT GAT AC-3'	38	296	93
02	BRCA2-20R/M13R	5'-CAG GAA ACA GCT ATG ACC ATG TTA AAT TCA AAG TCT CTA-3'	39		94
2 2	BRCA2-21F/M13F	5'-TGT AAA ACG ACG GCC AGT GGG TGT TTT ATG CTT GGT TCT-3'	39	304	95
2 2	BRCA2-21R/M13R	5'-CAG GAA ACA GCT ATG ACC CAT TTC AAC ATA TTC CTT CCT G-3'	40		%
3	BRCA2-22F-1A	5'-AAC CAC ACC CTT AAG ATG A-3'	19	453	26
23 62	BRCA2-22R-1A	5-GCA TTA GTA GTG GAT TTT GC-3'	50		86
1 6	BRCA2-23FII	5-TCA CTT CCA TTG CAT C-3'	16	290	66
23 23	BRCA2-23RII	5-TGC CAA CTG GTA GCT CC-3'	17		100
24	BRCA2-24 2F	5-TAC AGT TAG CAG CGA CAA AA-3'	50	373	101
7 7	BRCA2-24R/M13R	5-CAG GAA ACA GCT ATG ACC ATT TGC CAA CTG GTA GCT CC-3'	38		102
7 7	BRCA2-25F-7/23	5-GCT TTC GCC AAA TTC AGC TA-3'	20	427	103
25	BRCA2-25R-7/23		20		104
25	BRCA2/26-2F	5-AAT CAC TGA TAC TGG TTT TG-3'	20	530	105
36	BRCA2/26-2R	5'-TAT ACT TAC AGG AGC CAC AT-3'	20		106
27A	BRCA2-27AF-1A	6'-CTG TGT GTA ATA TTT GCG-3'	18	495	107
27.6	BRCA2-27AR/M13R	5-CAG GAA ACA GCT ATG ACG GCA AGT TCT TCG TCA GCT ATT G-3'	40		108
27B	BRCA2-27BF/M13F	3-TGT AAA AC3 ACG GCC AGT GAA TTC TCC TCA GAT GAC TCC A-3'	40	417	109
278	BRCA2-27BR/M13R	5'-CAG GAA ACA GCT ATG ACC TCT TTG CTC ATT GTG CAA CA-3'	38		110

TABLE III NORMAL PANEL TYPING

Position nt/codon	Nucleotide Change	Amino Acid Change	-	8	င	4	rC	Frequency
1093/289	<u>A</u> AT → <u>C</u> AT	Asn → His	A/A	AC	WA	A/A	A/C	A = .8 C = .2
1342/372	AAT → CAT	Asn → His	AC	A/A	AC	A/C	A/C	A = 0.6 C = 0.4
1593/455	TC <u>A</u> → TC <u>G</u>	Ser → Ser	AVA	A/A	AVA	A/A	A/G	A = 0.9 G = 0.1
2457/743	CAĪ→CAC	His → His	1/1	C/T	1/1	1/1	СЛ	T = 0.8 C = 0.2
2908/894	GTA → <u>A</u> TA	Val → Ile	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1
3199/991	<u>A</u> AC → <u>G</u> AC	Asn → Asp	A/A	A/G	A/A	A/A	A/G	A = 0.8 G = 0.2

TABLE III NORMAL PANEL TYPING

Position	Nucleotide Change	Amino Acid Change	~	2	က	4	Ŋ	Frequency
3624/1132	AA <u>A</u> → AA <u>G</u>	Lys → Lys	A/A	A/G	A/A	A/G	WA	A = 0.8 G = 0.2
4035/1269	GTI → GT <u>C</u>	Val → Val	C/T	T/T	1/1	1/1	1/1	T = 0.9 C = 0.1
7470/2414	TC <u>A</u> → TC <u>G</u>	Ser → Ser	AA A	A/G	A/A	AG	A/A	A = 0.8 G = 0.2
9079/2951	GCC → ACC	Ala → Thr	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1

EXAMPLE 2

5

20

25

<u>Determination Of A Normal Individual Using BRCA2^(OMI 1-5) and The Ten</u> <u>Polymorphisms For Reference</u>

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- identifying individuals having a normal BRCA2 gene;
- avoiding misinterpretation of normal polymorphisms found in the normal population.

Sequencing was carried out as in EXAMPLE 1 using a blood sample from the patient in question. However, the BRCA2 sequences were used for reference and any polymorphic sites seen in the patient were compared to the nucleic acid sequences listed above for normal codons at each polymorphic site. A normal sample is one which is comparable to the BRCA2 sequences and contains only minor variations which occur at minor polymorphic sites. The allelic variations which occur at each of the polymorphic sites are paired here for reference.

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- <u>C</u>AT (His) and <u>A</u>AT (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)
- CAT (His) and CAC (His) at position 2457 (codon 743)
- GTA (Val) and ATA (Ile) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GT<u>T</u> (Val) and GT<u>C</u> (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

The availability of these polymorphic pairs provides added assurance that one skilled in the art can correctly interpret the polymorphic variations without mistaking a normal variation for a mutation.

All exons of the BRCA2 gene are subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et

al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer[®] cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

10 1. Polymerase Chain Reaction (PCR) Amplification

5

15

20

25

30

The PCR primers used to amplify a patient's sample BRCA2 gene are listed in TABLE II. The primers were synthesized on a DNA/RNA Synthesizer Model 394[®]. Thirty-five cycles are of amplification are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick PCR purification kits (Qiagen cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD_{260} on a Beckman DU 650 spectrophotometer.

2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer[®] cat# 401628). DNA sequencing is performed in both rorward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator[®]" purchased through ABI. The BRCA2^(ornl 1-5) sequences were entered sequentially into the Sequence Navigator software as the standards for comparison. The Sequence Navigator software compares the patient sample sequence to each BRCA2 ^(ornl 1-5) standard, base by base. The Sequence Navigator highlights all differences between the standards (omi 1-5) and the patient's sample sequence.



A first technologist checks the computerized results by comparing visually the BRCA2 (omi 1-5) standards against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known normal polymorphisms for position and base change.

3. Results

5

10

15 nucleotide positions: 1093 (A/C), 1342 (A/C), 1593 (A/G), 2457 (C/T), 2908 (A/G), 3199 (A/G) and 9079 (A/G). In addition, this changes five amino acids in the polypeptide product: Asn to His at codon 289, Asn to His at codon 372, Val to lle at codon 894, Asn to Asp at codon 991, and Ala to Thr at codon 2951. The question arises whether any or all of these changes have significance to the patient.

20 Comparison of the patient's results to the BRCA (omi 1-5) haplotypes demonstrates that it matches one of the BRCA2 omi standards (#5), and thus the patient sample is interpreted as carrying a normal gene sequence without causing any elevation in their risk status for breast cancer.

25 **EXAMPLE 3**

30

DETERMINING THE PRESENCE OF A MUTATION IN EXON 11 OF THE BRCA2 GENE USING BRCA2(omi1-5)

A person skilled in the art of genetic susceptibility testing will find the present invention useful for determining the presence of a known or previously unknown mutation in the BRCA2 gene. A list of mutations of BRCA2 is publicly available in the Breast Cancer Information Core at http://www.nchgr.nih.gov/dir/lab_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. et al. Nature Genetics 11:238, (1995).

In this example, a mutation in exon 11 is characterized by amplifying the region of the mutation with a primer set which amplifies the region of the mutation. Sequencing was carried out as in Example 1 using a blood sample from the patient in question. Specifically, exon 11 of the BRCA2 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer® cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

15

20

25

10

5

1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl₂), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA2-11Q-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA2-11Q-R, 10 micromolar solution), and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify segment Q of exon 11 (where the mutation 6174delT is found) are as follows:

BRCA2-11Q-F: 5'- ACG' AAA' ATT' ATG' GCA' GGT' TGT-3'

BRCA2-11Q-R: 5'- CTT' GTC' TTG' CGT' TTT' GTA' ATG-3'

30

The primers are synthesized on an DNA/RNA Synthesizer Model 394[®]. Thirty-five cycles are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the

first cycle in which the denaturing time is increased to 5 minutes, and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick[®] PCR purification kits (Qiagen[®], cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD₂₆₀ on a Beckman DU 650 spectrophotometer.

2. <u>Dideoxy Sequence Analysis</u>

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer[®] cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator[®]" purchased through ABI. The BRCA2^(omi 1-5) sequence is entered into the Sequence Navigator software as the Standard for comparison. The Sequence Navigator software compares the sample sequence to the BRCA2^(omi) standard, base by base. The Sequence Navigator highlights all differences between the BRCA2^(omi) normal DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2^(omi 1-5) standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a sequence variation is compared to known normal polymorphisms for position and base change. The ten frequent polymorphisms which occur in BRCA2 are:

30

25

5

10

15

20

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- <u>CAT</u> (His) and <u>AAT</u> (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)

- CAT (His) and CAC (His) at position 2457 (codon 743)
- GTA (Val) and ΔTA (IIe) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GTT (Val) and GTC (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

10 3. Results

5

15

20

30

Using the above PCR amplification and standard fluorescent sequencing technology, the 6174delT mutation may be found. Mutations are noted by the length of non-matching sequence variation. Such a lengthy mismatch pattern occurs with deletions and insertions. This mutation is named in accordance with the suggested nomenclature for naming mutations, Beaudet, A *et al.*, Human Mutation 2:245-248, (1993). The 6174delT mutation at codon 1982 of the BRCA2 gene lies in segment "Q" of exon 11. The DNA sequence results demonstrate the presence of a one base pair deletion of a T at nucleotide 6174 of the BRCA2^(orni 1-5) sequences. This mutation interrupts the normal reading frame of the BRCA2 transcript, resulting in the appearance of an in-frame terminator (TAG) at codon position 2003. This mutation is, therefore, predicted to result in a truncated, and most likely, non-functional protein.

EXAMPLE 4

25 GENERATION OF MONOCLONAL AND POLYCLONAL ANTIBODIES USING GST-BRCA2 FUSION PROTEIN AS AN IMMUNOGEN

DNA primers are used to amplify a fragment of BRCA2 using PCR technology. The product is then digested with suitable restriction enzymes and fused in frame with the gene encoding glutathione S-transferase (GST) in *Escherichia coli* using GST expression vector pGEX (Pharmacia Biotech Inc.) The expression of the fusion protein is induced by the addition of isopropyl-β-thiogalactopyranoside. The bacteria are then lysed and the overexpressed fusion protein is purified with glutathione-sepharose beads. The fusion protein is then verified by SDS/PAGE gel and N-terminus protein sequencing. The purified protein

is used to immunize rabbits according to standard procedures described in Harlow & Lane (1988). Polycolonal antibody is collected from the serum several weeks after and purified using known methods in the art. Monoclonal antibodies against all or fragments of BRCA2 protein, polypeptides, or functional equivalents are obtained using hybridoma technology, see also Harlow & Lane (1988). The BRCA2 protein or polypeptide is coupled to the carrier keyhole limpet hemocyanin in the presence of glutaraldehyde. The conjugated immunogen is mixed with an adjuvant and injected into rabbits. Spleens from antibody-containing rabbits are removed. The B-cells isolated from spleen are fused to myeloma cells using polyethylene glycol (PEG) to promote fusion. The hybrids between the myeloma and B-cells are selected and screened for the production of antibodies to immunogen BRCA2 protein or polypeptide. Positive cells are recloned to generate monoclonal antibodies.

EXAMPLE 5

5

10

15

20

25

30

<u>DETECTION OF BRCA2 EXPRESSION IN HUMAN TISSUES AND CELL LINES</u>

The expression of BRCA2 in human tissues is determined using Northern blot analysis. Human tissues include those from pancreas, testis, prostate, ovary, breast, small intestine, and colon are obtained from Clontech Laboratories, Inc., Palo Alto, CA. The poly(A)+ mRNA Northern blots from different human tissues is hybridized to BRCA2 cDNA probes according to manufacture protocol. The expression level is further conformed by RT-PCR using oligo-d(T) as a primer and other suitable primers.

For Northern Blot analysis of cancer cell lines, the human ovarian cancer cell line SKOV-3 and the human breast cancer cell line MCF-7 are obtained from the American Type Culture Collection. Total RNA is prepared by lysing cell in the presence of guanidinium isocyanate. Poly(A)* mRNA is isolated using the PolyATract mRNA isolation system from Promega, Madison, WI. The isolated RNA is then electrophoresed under denaturing conditions and transferred to Nylon membrane. The probe used for Northern blot is a fragment of BRCA2 sequence obtained by PCR amplification. The probes are labeled with [α -32P] dCTP using a random-primed labeling kit (Amersham Life Science, Arlington Heights, IL).

EXAMPLE 6

5

10

15

20

25

30

EXPRESSION OF THE BRCA2 PROTEIN

The whole-cell extracts of BRCA2 transfected cells are subjected to immunoprecipitation and immunoblotting to determine the BRCA2 protein level. The BRCA2 protein or polypeptide is immunoprecipitated using anti-BRCA2 antibodies prepared according to Example 4. Samples are then fractionated using SDS/PAGE gel and transferred to nitrocellulose. Western blot of the BRCA2 protein or polypeptide is performed with the indicated antibodies. Antibody reaction is revealed using enhanced chemiluminescence reagents (Dupont New England Nuclear, Boston, MA).

EXAMPLE 7

USE OF THE BRCA2 (oml1-5) GENE THERAPY

The growth of ovarian or breast cancer may be arrested by increasing the expression of the BRCA2 gene where inadequate expression of that gene is responsible for hereditary ovarian or breast cancer. Gene therapy may be performed on a patient to reduce the size of a tumor. The LXSN vector may be transformed with a BRCA2^(omi1-5) coding sequence as presented SEQ ID NO:4, 6, 8, 10, or 12 or a fragment thereof.

Vector

The LXSN vector is transformed with a fragment of the wildtype BRCA2^(omi1-5) coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, or 12. The LXSN-BRCA2^(omi1-5) retroviral expression vector is constructed by cloning a *Sal* I linkered BRCA2^(omi1-5) cDNA or fragments thereof into the *Xho* I site of the vector LXSN. Constructs are confirmed by DNA sequencing. See Holt et al., *Nature Genetics* 12: 298-302 (1996). Retroviral vectors are manufactured from viral producer cells using serum free and phenol-red free conditions and tested for sterility, absence of specific pathogens, and absence of replication-competent retrovirus by standard assays. Retrovirus is stored frozen in aliquots which have been tested.

Patients receive a complete physical exam, blood, and urine tests to determine overall health. They may also have a chest X-ray, electrocardiogram, and appropriate radiologic procedures to assess tumor stage.

Patients with metastatic ovarian cancer are treated with retroviral gene therapy by infusion of recombinant LXSN-BRCA2^(omi1-5) retroviral vectors into peritoneal sites containing tumor, between 10⁹ and 10¹⁰ viral particles per dose. Blood samples are drawn each day and tested for the presence of retroviral vector by sensitive polymerase chain reaction (PCR)-based assays. The fluid which is removed is analyzed to determine:

- 1. The percentage of cancer cells which are taking up the recombinant LXSN-BRCA2^(omi1-5) retroviral vector combination. Successful transfer of BRCA1 gene into cancer cells has been shown by both RT-PCR analysis and *in situ* hybridization. RT-PCR is performed with by the method of Thompson et al., *Nature Genetics* 9: 444-450 (1995), using primers derived from a BRCA2^(omi1-5) coding sequence as in SEQ ID NO:4, 6, 8, 10, or 12 or fragments thereof. Cell lysates are prepared and immunoblotting is performed by the method of Jensen *et al.*, *Nature Genetics* 12: 303-308 (1996) and Jensen *et al.*, *Biochemistry* 31: 10887-10892 (1992).
- 2. Presence of programmed cell death using APOTAG® in situ apoptosis detection kit (ONCOR, INC., Gaithersburg, Maryland) and DNA analysis.
- Measurement of BRCA2 gene expression by slide immunofluorescence or
 Western blot.

Patients with measurable disease are also evaluated for a clinical response to LXSN-BRCA2^(omi1-5) especially those that do not undergo a palliative intervention immediately after retroviral vector therapy. Fluid cytology, abdominal girth, CT scans of the abdomen, and local symptoms are followed.

25

30

5

10

15

For other sites of disease, conventional response criteria are used as follows:

- 1. Complete Response (CR), complete disappearance of all measurable lesions and of all signs and symptoms of disease for at least 4 weeks.
- 2. Partial Response (PR), decrease of at least 50% of the sum of the products of the 2 largest perpendicular diameters of all measurable lesions as determined by 2 observations not less than 4 weeks apart. To be considered a PR, no new lesions should have appeared during this period and none should have increased in size.
 - 3. Stable Disease, less than 25% change in tumor volume from previous evaluations.

4. Progressive Disease, greater than 25% increase in tumor measurements from prior evaluations. The number of doses depends upon the response to treatment.

EXAMPLE 8

5

10

15

20

25

PROTEIN REPLACEMENT THERAPY

Therapeutically elevated level of functional BRCA2 protein may alleviate the absence or reduced endogenous BRCA2 tumor suppressing activity. Breast or ovarian cancer is treated by the administration of a therapeutically effective amount of the BRCA2 protein, a polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Clinically effective delivery method is applied either locally at the site of the tumor or systemically to reach other metastasized locations with known protocols in the art. These protocols may employ the methods of direct injection into a tumor or diffusion using time release capsule. A therapeutically effective dosage is determined by one of skill in the art.

Breast or ovarian cancer may be prevented by the administration of a prophylactically effective amount of the BRCA2 protein, polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Individuals with known risk for breast or ovarian cancer are subjected to protein replacement therapy to prevent tumorigenesis or to decrease the risk of cancer. Elevated risk for breast and ovarian cancer includes factors such as carriers of one or more known BRCA1 and BRCA2 mutations, late child bearing, early onset of menstrual period, late occurrence of menopause, and certain high risk dietary habits. Clinically effective delivery method is used with known protocols in the art, such as administration into peritoneal cavity, or using an implantable time release capsule. A prophylactically effective dosage is determined by one of skill in the art.

TABLE OF REFERENCES

- 1. Sanger, F., et al., J. Mol. Biol. 42:1617, (1980).
- 30 2. Beaucage, et al., Tetrahedron Letters 22:1859-1862, (1981).
 - 3. Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, (1982).
 - 4. Conner, et al., Proc. Natl. Acad. Sci. U.S.A. 80:278, (1983)
 - 5. Saiki, et.al.., *Bio/Technology* 3:1008-1012, (1985)

- 6. Landgren, et al., Science 241:1007,(1988)
- 7. Landgren, et al., Science 242:229-237, (1988).
- 8. PCR. A Practical Approach, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, (1992).
- 9. Easton et al., American Journal of Human Genetics 52:678-701, (1993).
- 10. Patent No. 4,458,066.

5

20

- 11. Rowell, S., et al., American Journal of Human Genetics 55:861-865, (1994)
- 12. Miki, Y. et al., Science 266:66-71, (1994).
- 10 13. Wooster, R. et al., Science <u>265</u>:2088-2090, (1994).
 - 14. Wooster, R. et al., Nature 378:789-792, (1095).
 - 15. Beaudet, A et al., Human Mutation 2:245-248, (1993).
 - 16. Friend, S. et al. Nature Genetics 11:238. (1995).
 - 17. Teng et al, Nature Genetics 13: 241-244 (1996).
- 15 18. Couch et al, Nature Genetics 13: 123-125 (1996).
 - 19. Tartigan et al, Nature Genetics 12: 333-337 (1996).
 - 20. Phelan et al, Nature Genetics 13: 120-122 (1996).
 - 21. Schubert et al, American Journal of Human Genetics 60: 1031-1040 (1996).
 - 22. Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989).
 - 23. Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998).
 - 24. Zhang et al., Cell <u>92</u>:433-436 (1998).
 - 25. Sharan et al., Nature 386:804-810 (1997).
 - 26. Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988).
- 25 27. Crooke, Annu. Rev. Pharmacol. Toxicol. 32:329-376 (1992)
 - 28. Robinson-Benion and Holt, Methods Enzymol. 254:363-375 (1995).
 - 29. Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988.
- 30. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486,
 DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993.
 - 31. Holt et al., Nature Genetics 12: 298-302 (1996).
 - 32. Thompson et al., Nature Genetics 9: 444-450 (1995).
 - 33. Jensen et al., Nature Genetics 12: 303-308 (1996)
 - 34. Jensen et al., Biochemistry 31: 10887-10892 (1992).

Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

5

SEQUENCE LISTING

5	(1) GENERAL INFORMATION
	(i) APPLICANT: Murphy, Patricia
	(1) APPLICANT: Mulphy, Facticia White, Marga
10	Rabin, Mark
	Olson, Sheri
	Yoshikawa, Matthew Jackson, Geoffrey
	Eskanderi, Tara
15	Schryer, Brenda .
	Park, Michael
	(ii) TITLE OF THE INVENTION: NOVEL CODING SEQUENCE HAPLOTYPES
20	OF THE HUMAN BRCA2 GENE
20	(iii) NUMBER OF SEQUENCES: 111
	(iv) CORRESPONDENCE ADDRESS:
	(A) ADDRESSEE: Howrey & Simon
25	(B) STREET: 1299 Pennsylvania Avenue N.W.
	(C) CITY: Washington (D) STATE: DC
	(E) COUNTRY: USA
	(F) ZIP: 20004
30	(v) COMPUTER READABLE FORM:
	(A) MEDIUM TYPE: Diskette
	(B) COMPUTER: IBM Compatible
35	(C) OPERATING SYSTEM: DOS(D) SOFTWARE: FastSEQ for Windows Version 2.0
35	
	(vi) CURRENT APPLICATION DATA:
	(A) APPLICATION NUMBER: (B) FILING DATE:
40	(C) CLASSIFICATION:
	A LIA MARCO ADDITIONAL DAMA
	<pre>(vii) PRIOR APPLICATION DATA: (A) APPLICATION NUMBER:</pre>
	(B) FILING DATE:
45	
	A LALL A CONTRACTOR TAYON TO THE TOWN
	<pre>(viii) ATTORNEY/AGENT INFORMATION: (A) NAME: Halluin, Albert P</pre>
50	(B) REGISTRATION NUMBER: 25,227
	(C) REFERENCE/DOCKET NUMBER: 5371.31.US02
	(ix) TELECOMMUNICATION INFORMATION:
	(A) TELEPHONE: 650-463-8109
55	(B) TELEFAX: 650-463-8400 (C) TELEX:
	(C) LEADER.
	(2) INFORMATION FOR SEQ ID NO:1:
60	(i) GEOLUTION GUADACTEDICTICS.
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 50 base pairs

5	(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear	
5	(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:	
10	(A) NAME/KEY: exon(B) LOCATION: 150(D) OTHER INFORMATION: Exon 5	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
15	TCCTGTTGTT CTACAATGTA CACATGTAAC ACCACAAAGA GATAAGTCAG	50
	(2) INFORMATION FOR SEQ ID NO:2:	
20	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 182 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear 	
25	(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:	
30	(A) NAME/KEY: exon (B) LOCATION: 1182 (D) OTHER INFORMATION: Exon 15	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:	
35	ATTTAATTAC AAGTCTTCAG AATGCCAGAG ATATACAGGA TATGCGAATT AAGAAGAAAC AAAGGCAACG CGTCTTTCCA CAGCCAGGCA GTCTGTATCT TGCAAAAACA TCCACTCTGC CTCGAATCTC TCTGAAAGCA GCAGTAGGAG GCCAAGTTCC CTCTGCGTGT TCTCATAAAC AG	60 120 180 182
40	(2) INFORMATION FOR SEQ ID NO:3:	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 188 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double	
45	(D) TOFOLOGY: linear	
	<pre>(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:</pre>	
50	(A) NAME/KEY: exon(B) LOCATION: 1188(D) OTHER INFORMATION: Exon 16	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:	
33	CTGTATACGT ATGGCGTTTC TAAACATTGC ATAAAAATTA ACAGCAAAAA TGCAGAGTCT TTTCAGTTTC ACACTGAAGA TTATTTTGGT AAGGAAAGTT TATGGACTGG AAAAGGAATA CAGTTGGCTG ATGGTGGATG GCTCATACCC TCCAATGATG GAAAGGCTGG AAAAGAAGAA TTTTTATAG	60 120 180 188
60	(2) INFORMATION FOR SEQ ID NO:4:	

5		((A) I (B) I (C) S	ENGT TYPE : TRAN	TH: I nuc NDEDI	10485 cleid	TERI bas c aci : sin near	se pa ld	CS: airs								
10			(A) (B)	NAMI	RE: E/KE: ATIO	Y: Co N: 2:	29	g Se	queno 82 BRCA:		MI1)						
15		(x :	i) SI	EQUE	NCE I	DESC	RIPT	ION:	SEQ	ID I	NO : 4	:					
20	mama	CTGC	GC C'	TCGG	GTGT	C TT	TTGC: TTTT	GGCG GTCA	GTG GCT	GGTC FACT	GCC CCG	GCCG(GCCA)	GGAG. AAAA	AA G AG A ATG	ACTG ACTG CCT	GCGCC AGGGG CACCT ATT Ile	60 120 180 237
25	GGA '	TCC Ser 5	AAA Lys	GAG .	AGG Arg	Pro	ACA Thr 10	TTT Phe	TTT Phe	GAA Glu	Ile	TTT . Phe :	AAG Lys	ACA Thr	CGC Arg	TGC Cys	285
30	AAC Asn 20	Lys	Ala	Asp	Leu	Gly 25	Pro	Ile	Ser	Leu	Asn 30	Trp	Pne	GIU	GIU	35	333
2.5	TCT Ser	TCA Ser	GAA Glu	Ala	CCA Pro 40	CCC Pro	TAT Tyr	AAT Asn	TCT Ser	GAA Glu 45	CCT Pro	GCA Ala	GAA Glu	GAA Glu	TCT Ser 50	GAA Glu	381
35	CAT His	AAA Lys	AAC Asn	AAC Asn 55	AAT Asn	TAC Tyr	GAA Glu	CCA Pro	AAC Asn 60	CTA Leu	TTT Phe	AAA Lys	ACT Thr	CCA Pro 65	CAA Gln	AGG Arg	429
40	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	AAT Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
45	CAA Gln	GGG Gly 85	CTG Leu	ACT Thr	CTG Leu	CCG Pro	CTG Leu 90	TAC Tyr	CAA Gln	TCT Ser	CCT Pro	GTA Val 95	AAA Lys	GAA Glu	TTA Leu	GAT Asp	525
50	AAA Lys 100	TTC Phe	AAA Lys	TTA Leu	GAC Asp	TTA Leu 105	GGA Gly	AGG Arg	AAT Asn	GTT Val	CCC Pro 110	AAT Asn	AGT Ser	AGA Arg	CAT His	AAA Lys 115	573
	AGT Ser	CTT Leu	CGC Arg	ACA Thr	GTG Val 120	Lys	ACT Thr	AAA Lys	ATG Met	GAT Asp 125	CAA Gln	GCA Ala	GAT Asp	GAT Asp	GTT Val 130	TCC Ser	621
55	TGT Cys	CCA Pro	CTT Leu	CTA Leu 135	Asn	TCT Ser	TGT	CTT Leu	AGT Ser 140	GAA Glu	AGT Ser	CCT Pro	GTT Val	GTT Val 145	CTA Leu	CAA Gln	669
60	TGT Cys	ACA Thr	CAT His	Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	Cys	GGG Gly	AGT Ser	717

5	TTG Leu	TTT Phe 165	CAT His	ACA Thr	CCA . Pro	Lys	TTT Phe 170	GTG Val	AAG Lys	GGT Gly	CGT Arg	CAG Gln 175	ACA Thr	CCA Pro	AAA Lys	CAT His	765
	ATT Ile 180	TCT Ser	GAA Glu	AGT Ser	CTA Leu	GGA Gly 185	GCT Ala	GAG Glu	GTG Val	GAT Asp	CCT Pro 190	GAT Asp	ATG Met	TCT Ser	TGG Trp	TCA Ser 195	813
10	AGT Ser	TCT Ser	TTA Leu	GCT Ala	ACA Thr 200	CCA Pro	CCC Pro	ACC Thr	CTT Leu	AGT Ser 205	TCT Ser	ACT Thr	GTG Val	CTC Leu	ATA Ile 210	GTC Val	861
15	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
20	AAT Asn	GTG Val	AAA Lys 230	AGC Ser	TAT Tyr	TTT Phe	TCC Ser	AAT Asn 235	CAT His	GAT Asp	GAA Glu	AGT Ser	CTG Leu 240	AAG Lys	AAA Lys	AAT Asn	957
25	GAT Asp	AGA Arg 245	Phe	ATC Ile	GCT Ala	TCT Ser	GTG Val 250	ACA Thr	GAC Asp	AGT Ser	GAA Glu	AAC Asn 255	ACA Thr	AAT Asn	CAA Gln	AGA Arg	1005
	GAA Glu 260	GCT Ala	GCA Ala	AGT Ser	CAT His	GGA Gly 265	TTT Phe	GGA Gly	AAA Lys	ACA Thr	TCA Ser 270	GIY	AAT Asn	TCA Ser	TTT	AAA Lys 275	1053
30	GTA Val	AAT	AGC Ser	TGC Cys	AAA Lys 280	GAC Asp	CAC	ATT Ile	GGA Gly	AAG Lys 285	TCA Ser	ATG Met	CCA Pro	AAT Asn	GTC Val 290	CTA Leu	1101
35	GAA Glu	GAT Asp	GAA Glu	GTA Val 295	Tyr	GAA Glu	ACA Thr	GTT Val	GTA Val 300	Asp	ACC Thr	TCT Ser	GAA Glu	GAA Glu 305	Map	AGT Ser	1149
40	TTI Phe	TC#	TTA Leu	Cys	TTT Phe	TCT	Lys	TGT Cys 315	Arg	ACA Thr	AAA Lys	AAT Asn	CTA Leu 320	CAA Gln	AAA Lys	GTA Val	1197
45	AGA Arg	ACT Thi	r Sei	AAG Lys	ACT Thr	AGG	AAA Lys 330	Lys	ATT	TTC Phe	CAT His	GAA Glu 335	i Ala	AAC Asi	GCT Ala	GAT Asp	1245
	GA/ Glu 340	з Су	r GAJ s Gli	A AAA 1 Lys	A TCI	Lys	Ası	C CAA	A GTO	AAA L Lys	GA/ Glu 350	у гув	TAC Tyr	TC#	TTI Phe	GTA val	1293
50	TC' Se:	r GA r Gl	A GTO	G GAA	A CCA u Pro 360	Ası	GA' 1 As	r AC'	r GA:	r CCZ p Pro 365	o Lei	A GA? u Asj	r TCA p Sei	AA A	T GTA 1 Va. 37	A GCA l Ala 0	1341
55	CA Hi	T CA s Gl	G AA n Ly	G CCC s Pro	o Phe	r GAG	G AG u Se	T GG	A AGʻ y Se: 38	r As	C AA	A ATO	C TCC	C AAC Ly: 38	B GI	A GTT u Val	1389
60	GT Va	A CC	G TC O Se 39	r Le	G GCC u Ala	C TG a Cy	T GA s Gl	A TG u Tr 39	p Se	T CA r Gl	A CT n Le	A AC	C CT r Le 40	u se	A GG r Gl	T CTA y Leu	1437

	WO >	7/071															
-	Asn	Gly 405	Ala	CAG Gln	Met	Glu	Lys 410	Ile	Pro	Leu	Leu	415	116	ser	261	Cys	1485
5	Asp 420	Gln	Asn	ATT Ile	Ser	Glu 425	Lys	Asp	ьeu	ьеu	430	1111	GIU	ADII	2,5	435	1533
10	Lys	Lys	Asp	TTT Phe	Leu 440	Thr	ser	GIU	Asn	445	ьeu	PIO	Arg	110	450	501	1581
15	Leu	Pro	Lys	TCA Ser 455	Glu	Lys	Pro	Leu	Asn 460	Glu	GIU	Thr	Val	465	ASII	цуз	1629
20	Arg	Asp	470	GAG Glu	Gln	His	Leu	Glu 475	Ser	His	Tnr	Asp	480	116	ьеu	AIG	1677
٥٢	GTA Val	AAG Lys 485	Gln	GCA Ala	ATA Ile	TCT Ser	GGA Gly 490	ACT Thr	TCT Ser	CCA Pro	GTG Val	GCT Ala 495	TCT Ser	TCA Ser	TTT	CAG Gln	1725
25	Gly 500	Ile	Lys	AAG Lys	Ser	Ile 505	Phe	Arg	Ile	Arg	510	ser	PIO	Буз	GIU	515	1773
30	TTC Phe	AAT Asn	GCA Ala	AGT Ser	TTT Phe 520	TCA Ser	GGT Gly	CAT His	ATG Met	ACT Thr 525	Asp	CCA Pro	AAC Asn	TTT Phe	AAA Lys 530	٠,٠	1821
35	GAA Glu	ACT Thr	GAP	GCC Ala 535	Ser	GAA Glu	AGT Ser	GGA Gly	CTG Leu 540	Glu	ATA Ile	CAT His	ACT Thr	GTT Val 545	Cys	TCA Ser	1869
40	CAG Gln	AAC Lys	GA0 Glu 550	qaA ı	TCC Ser	TTA Leu	TGT Cys	CCA Pro 555	Asn	TTA Leu	ATT	GAT Asp	AAT Asn 560	Gry	AGC Ser	TGG Trp	1917
	CCP Pro	GC0 Ala 56	a Thi	C ACC	ACA Thr	Gln	Asr	TCT Ser	· vai	. Ala	і гег	, DAS	ASI	GCA	GGT Gly	TTA Leu	1965
45	AT# 116	e Se	C AC'	r TTC r Leu	AAA Lys	AAG Lys 585	Lys	A ACA	AAT Asr	T AAC	TT:	s TTE	TAT Tyr	GCT Ala	TATA 11e	CAT His 595	2013
50	GA' Asj	T GA p Gl	A AC u Th	A TC	TAT TY1 600	Lys	GGI Gly	A AAA y Lys	A AAI s Ly:	A ATA	e Pro	G AAI o Ly:	A GAC	CA/	A AAA 1 Lys 610	A TCA s Ser	2061
55	GA Gl	A CT u Le	A AT	T AAG e As: 61	n Cy	r TC	A GC	c cac a Gli	TT n Pho 62	e GI	A GC u Al	A AA' a Asi	r GCT n Ala	TT' a Pho 62	e Gr	A GCA u Ala	2109
60	CC Pr	A CI	T AC u Th	r Ph	T GC	A AA' a As:	r GC n Al	T GA' a As; 63	p Se	A GG r Gl	T TT y Le	A TT u Le	G CA' u Hi 64	s se	T TC ~ Se	T GTG r Val	2157
	A.A	A A	A AC	C TG	T TC	A CA	g aa	T GA	T TC	T GA	A GA	A CC	A AC	T TT	G TC	C TTA	2205

	Lys	Arg 645	Ser	Сув	Ser		Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr	Leu	Ser	Leu	
5	ACT Thr 660	AGC Ser	TCT Ser	TTT Phe	GGG Gly	ACA Thr 665	ATT Ile	CTG Leu	AGG Arg	AAA Lys	TGT Cys 670	TCT Ser	AGA Arg	AAT Asn	GAA Glu	ACA Thr 675	2253
10	TGT Cys	TCT Ser	AAT Asn	AAT Asn	ACA Thr 680	GTA Val	ATC Ile	TCT Ser	CAG Gln	GAT Asp 685	CTT Leu	GAT Asp	TAT Tyr	AAA Lys	GAA Glu 690	GCA Ala	2301
15	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
20	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
25	CCA Pro 740	Val	CAA Gln	CAT His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
30	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
35	ACT Thr	CCT Pro	ACT Thr	TCC Ser 775	AAG Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	TCT Ser	AGA Arg	2589
	GGC Gly	Lys	GAA Glu 790	Ser	TAC	AAA Lys	ATG Met	TCA Ser 795	GAC Asp	AAG Lys	CTC Leu	AAA Lys	GGT Gly 800	AAC Asn	AAT Asn	TAT Tyr	2637
40	GAP Glu	TCI Ser 805	qaA :	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	Lys	AAT Asn	ATT	CCC	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
45	GAT Asp 820	Val	A TGI L Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT	Lys	AAC Asn 830	Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
50	CC.	GA/	A AAA	A TAC	ATG Met	Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	Ser	AGA Arg	AAG Lys	GTA Val	Gln 850	TTC Phe	2781
55	AA As:	c CAl	A AA(n Asi	C ACA	Asn	CTA Leu	AGA	GTA Val	ATC 116	Glr	A AAA	AAT BASN	CAA Gln	GAA Glu 865	GIU	ACT Thr	2829
	AC' Th	T TC	A AT	e Se	A AAA	ATA ile	ACT	GTC Val	Ası	r CC#	A GAG	TCT Ser	GAP Glu 880	GII	CTI Lei	TTC Phe	2877
60	TC Se	A GA r As	C AA' p As	T GA	G AAT u Asi	AA 1 a Asi	TT:	r GTO	C TT(C CAM	A GTA	A GCT	CAA T	GAI	A AGO	G AAT J Asn	2925

PCT/US98/16905

WO 99/09164 AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TIG ACT Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile GAT AAA ATA CCA GAA AAA AAT AL. GAT TAC ATG AAC AAA TGG GCA GGA Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys Trp Ala Gly CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala TGT GTT GAA ATT G.A AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser GTA GTT GTT TCT GTT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu Ser Gly Ser

CAG TTT GAA TTT ACT CAG TTT AGA AAA CCA AGC TAC ATA TTG CAG AAG

Gin Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys

5	AGT ACA TTT GAA GTG CCT GAA AAC CAG AIG ACT ALC THE Leu Lys Thr Thr Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155	3693
	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
10	TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
15	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
20	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
25	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
30	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
35	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
40	AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
45	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310	4173
	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
50	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
55	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys lle His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1360	4317
60	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 1370 1375	4365

	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
5	Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
10	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
15	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
20	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605
o.=	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475	4653
25	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
30	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
35	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
40	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
4.5	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1545 1550 1555	4893
45	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570	4941
50	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
55	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
60	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG	5133

	Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	
5	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
10	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
15	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
20	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
25	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
30	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
35	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
40	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
45	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
50	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
55	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
_	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glv Thr Ile 1845 1850 1855	5805
60	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys	5853

	1860	1865		1870	1875
5	GAA AAC Glu Asn	AAC GAG AAT AAA Asn Glu Asn Lys 1880	TCA AAA ATT TGC Ser Lys Ile Cys 1885	Gln Thr Lys I	TT ATG GCA 5901 The Met Ala 1890
10	GGT TGT Gly Cys	TAC GAG GCA TTG Tyr Glu Ala Leu 1895	GAT GAT TCA GAG Asp Asp Ser Glu 1900	Asp Ile Leu F	CAT AAC TCT 5949 His Asn Ser 905
15	Leu Asp	AAT GAT GAA TGT Asn Asp Glu Cys 910	AGC ACG CAT TCA Ser Thr His Ser 1915	A CAT AAG GTT T His Lys Val I 1920	TTT GCT GAC 5997 Phe Ala Asp
15	ATT CAG Ile Gln 1925	AGT GAA GAA ATT Ser Glu Glu Ile	TTA CAA CAT AAC Leu Gln His Ası 1930	C CAA AAT ATG 7 n Gln Asn Met 9 1935	CCT GGA TTG 6045 Ser Gly Leu
20	GAG AAA Glu Lys 1940	GTT TCT AAA ATA Val Ser Lys Ile 1945	TCA CCT TGT GAT Ser Pro Cys Asp	r GTT AGT TTG () Val Ser Leu (1950	GAA ACT TCA 6093 Glu Thr Ser 1955
25	GAT ATA Asp Ile	TGT AAA TGT AGT Cys Lys Cys Ser 1960	ATA GGG AAG CTT Ile Gly Lys Lev 1965	ı His Lys Ser \	STC TCA TCT 6141 Val Ser Ser 1970
30	GCA AAT Ala Asn	ACT TGT GGG ATT Thr Cys Gly Ile 1975	TTT AGC ACA GCA Phe Ser Thr Ala 1980	a Ser Gly Lys :	TCT GTC CAG 6189 Ser Val Gln 985
	Val Ser	GAT GCT TCA TTA Asp Ala Ser Leu 1990	CAA AAC GCA AGA Gln Asn Ala Arg 1995	A CAA GTG TTT 1 g Gln Val Phe 1 2000	TCT GAA ATA 6237 Ser Glu Ile
35	GAA GAT Glu Asp 2005	AGT ACC AAG CAA Ser Thr Lys Gln	GTC TTT TCC AA Val Phe Ser Ly: 2010	A GTA TTG TTT A s Val Leu Phe 1 2015	AAA AGT AAC 6285 Lys Ser Asn
40	GAA CAT Glu His 2020	TCA GAC CAG CTC Ser Asp Gln Leu 2025	Thr Arg Glu Glu	A AAT ACT GCT A A Asn Thr Ala : 2030	ATA CGT ACT 6333 Ile Arg Thr 2035
45	CCA GAA Pro Glu	CAT TTA ATA TCC His Leu Ile Ser 2040	CAA AAA GGC TT Gln Lys Gly Pho 204	e Ser Tyr Asn	GTG GTA AAT 6381 Val Val Asn 2050
50	TCA TCT Ser Ser	GCT TTC TCT GGA Ala Phe Ser Gly 2055	TTT AGT ACA GC. Phe Ser Thr Al	a Ser Gly Lys	CAA GTT TCC 6429 Gln Val Ser 065
	Ile Leu	GAA AGT TCC TTA Glu Ser Ser Leu 2070	CAC AAA GTT AA His Lys Val Ly 2075	G GGA GTG TTA s Gly Val Leu 2080	GAG GAA TTT 6477 Glu Glu Phe
55	GAT TTA Asp Leu 2085	ATC AGA ACT GAG Ile Arg Thr Glu	CAT AGT CTT CA His Ser Leu Hi 2090	C TAT TCA CCT s Tyr Ser Pro 2095	ACG TCT AGA 6525 Thr Ser Arg
60	CAA AAT Gln Asn 2100	GTA TCA AAA ATA Val Ser Lys Ile 2105	Leu Pro Arg Va	T GAT AAG AGA 1 Asp Lys Arg 2110	AAC CCA GAG 6573 Asn Pro Glu 2115

5	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
1.0	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
10	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160	6717
15	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
20	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
25	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
30	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
30	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
35	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
40	GAG GAA ATG GTT T ^G TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	7053
45	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
50	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2305 2300 2305	7149
50	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
55	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
60	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2345 2350 2355	7293

_	TTT CTG TCT AAA T Phe Leu Ser Lys S 23	CT CAT TTG TAT GAA er His Leu Tyr Glu 60	CAT CTG ACT TTG (His Leu Thr Leu (2365	GAA AAA TCT 7341 Glu Lys Ser 2370
5	TCA AGC AAT TTA G Ser Ser Asn Leu A 2375	CA GTT TCA GGA CAT la Val Ser Gly His 2380	Pro Phe Tyr Gln	GTT TCT GCT 7389 Val Ser Ala 385
10	ACA AGA AAT GAA A Thr Arg Asn Glu L 2390	AA ATG AGA CAC TTG ys Met Arg His Leu 2395	ATT ACT ACA GGC A Ile Thr Thr Gly A 2400	AGA CCA ACC 7437 Arg Pro Thr
15	Lys Val Phe Val P 2405	CA CCT TTT AAA ACT ro Pro Phe Lys Thr 2410	Lys Ser His Phe 1 2415	His Arg Val
20	Glu Gln Cys Val A 2420	GG AAT ATT AAC TTG rg Asn Ile Asn Leu 2425	Glu Glu Asn Arg (2430	Gln Lys Gln 2435
25	AAC ATT GAT GGA C Asn Ile Asp Gly H 24	AT GGC TCT GAT GAT is Gly Ser Asp Asp 40	AGT AAA AAT AAG Ser Lys Asn Lys 2445	ATT AAT GAC 7581 Ile Asn Asp 2450
23	AAT GAG ATT CAT C Asn Glu Ile His G 2455	AG TTT AAC AAA AAC In Phe Asn Lys Asn 2460	Asn Ser Asn Gln	GCA GCA GCT 7629 Ala Ala Ala 465
30	GTA ACT TTC ACA A Val Thr Phe Thr L 2470	AG TGT GAA GAA GAA ys Cys Glu Glu Glu 2475	CCT TTA GAT TTA A Pro Leu Asp Leu 2480	ATT ACA AGT 7677 Ile Thr Ser
35	CTT CAG AAT GCC A Leu Gln Asn Ala A 2485	GA GAT ATA CAG GAT arg Asp Ile Gln Asp 2490	ATG CGA ATT AAG A Met Arg Ile Lys : 2495	AAG AAA CAA 7725 Lys Lys Gln
40	AGG CAA CGC GTC T Arg Gln Arg Val P 2500	TT CCA CAG CCA GGC The Pro Gln Pro Gly 2505	AGT CTG TAT CTT (Ser Leu Tyr Leu . 2510	GCA AAA ACA 7773 Ala Lys Thr 2515
45	Ser Thr Leu Pro A	GA ATC TCT CTG AAA Arg Ile Ser Leu Lys 320	GCA GCA GTA GGA G Ala Ala Val Gly G 2525	GGC CAA GTT 7821 Gly Gln Val 2530
43	CCC TCT GCG TGT T Pro Ser Ala Cys S 2535	CT CAT AAA CAG CTG Ser His Lys Gln Leu 2540	Tyr Thr Tyr Gly	GTT TCT AAA 7869 Val Ser Lys 545
50	CAT TGC ATA AAA A His Cys Ile Lys I 2550	ATT AAC AGC AAA AAT le Asn Ser Lys Asn 2555	GCA GAG TCT TTT Ala Glu Ser Phe 2560	CAG TTT CAC 7917 Gln Phe His
			משא שכם אכש ככא	AAA GGA ATA 7965
55	ACT GAA GAT TAT T Thr Glu Asp Tyr F 2565	Phe Gly Lys Glu Ser 2570	Leu Trp Thr Gly 2575	Lys Gly Ile
55 60	Thr Glu Asp Tyr F 2565 CAG TTG GCT GAT G	Phe Gly Lys Glu Ser	Leu Trp Thr Gly 2575 . CCC TCC AAT GAT	Lys Gly Ile GGA AAG GCT 8013

	Gly Ly	s Glu		Phe 2600	Tyr	Arg	Ala		Cys 2605	Asp	Thr	Pro	_	Val 2610	Asp	
5	CCA AA Pro Ly	s Leu					Trp					Tyr				8109
10	ATA TO					Met					Pro					8157
15	AAT AG Asn Ar 264	g Cys			Pro					Leu						8205
20	TAT GA Tyr As 2660			Ile					Ārg					Lys		8253
20	ATG GA Met Gl		Asp					Lys					Сув			8301
25	GAC AT Asp Il	e Ile					Asn					Ser				8349
30	ACT AG Thr Se					Gln					Ile					8397
35	GGG TG Gly Tr 272	p Tyr			Lys					Pro						8445
40	TTA AA Leu Ly 2740			Arg					Gln					His		8493
40	GCA GA Ala Gl		Val					Ala					Glu			8541
45	GAA TC Glu Se	r Leu					Ser					Arg				8589
50																8637
	TGG TA Trp Ty	T ACC r Thr 2790				Phe					Arg					
55		r Thr 2790 A TCA u Ser	Lys	Leu	Gly TTC Phe	Phe 2 AGT	Phe 795 GAT	Pro GGA	Asp GGA	Pro AAT Asn	Arg 2 GTT	Pro 800 GGT	Phe TGT	Pro GTT	Leu GAT	8685
55	Trp Ty CCC TT Pro Le	r Thr 2790 A TCA u Ser 5	TCG Ser	CTT Leu AGA Arg	Gly TTC Phe 2 GCA	AGT Ser 2810	Phe 2795 GAT Asp	Pro GGA Gly ATA	Asp GGA Gly CAG Gln	Pro AAT Asn 2 TGG	Arg GTT Val 815 ATG	Pro 800 GGT Gly	Phe TGT Cys AAG	Pro GTT Val ACA Thr	Leu GAT Asp TCA	8685 8733

2840 2845 2850

		2840		2845	2850	
5			Ala Gln Gln		GAA GCC TTA TTC Glu Ala Leu Phe 2865	
10		ln Glu Glu I			AAC ACA ACA AAA Asn Thr Thr Lys 2880	
15				Arg Gln Gln V	GTT CGT GCT TTG Val Arg Ala Leu 395	
		la Glu Leu '			GCA GCA GAC CCA Ala Ala Asp Pro	
20	TAC CTT G	AG GGT TAT 1 lu Gly Tyr 1 2920	TTC AGT GAA Phe Ser Glu	GAG CAG TTA A Glu Gln Leu A 2925	AGA GCC TTG AAT Arg Ala Leu Asn 2930	AAT 9021 Asn
25			Asn Asp Lys		CAG ATC CAG TTG Gln İle Gln Leu 2945	
30	ATT AGG AM Ile Arg Ly 299	ys Ala Met (GAA TCT GCT Glu Ser Ala 2955	GAA CAA AAG G Glu Gln Lys G	GAA CAA GGT TTA Glu Gln Gly Leu 2960	TCA 9117 Ser
35	AGG GAT G Arg Asp Va 2965	TC ACA ACC (GTG TGG AAG Val Trp Lys 2970	Leu Arg Ile V	GTA AGC TAT TCA Val Ser Tyr Ser 1975	AAA 9165 Lys
		ys Asp Ser V			CGT CCA TCA TCA Arg Pro Ser Ser	
40					AGA ATT TAT CAT Arg Ile Tyr His 3010	
45			Lys Ser Lys		GCT AAC ATA CAG Ala Asn Ile Gln 3025	
50		hr Lys Lys '			CCG GTT TCA GAT Pro Val Ser Asp 3040	
55				Arg Glu Pro I	CTT CAC TTC AGO Leu His Phe Ser 055	
		sp Pro Asp			BAG GTG GAC CTA Blu Val Asp Leu	
60					CTT GCC CCT TTC Leu Ala Pro Phe 3090	Val

5		p Glu Cys Tyr Asr	TTA CTG GCA ATA A Leu Leu Ala Ile L 3100	
10	GAC CTT AAT GA Asp Leu Asn Gl 3110	G GAC ATT ATT AAG u Asp Ile Ile Lys 3115	CCT CAT ATG TTA A Pro His Met Leu I 31	le Ala Ala Ser
10	AAC CTC CAG TG Asn Leu Gln Tr 3125	G CGA CCA GAA TCC p Arg Pro Glu Ser 3130	AAA TCA GGC CTT C Lys Ser Gly Leu L 3135	TT ACT TTA TTT 9645 eu Thr Leu Phe
15			GCT AGT CCA AAA G Ala Ser Pro Lys G 3150	
20			AAT ACT GTT GAG A Asn Thr Val Glu A 3165	
25		u Ala Glu Asn Lys	CTT ATG CAT ATA C Leu Met His Ile L 3180	
2.0			AAA GAC TGT ACT T Lys Asp Cys Thr S 32	er Gly Pro Tyr
30			GGA AAC AAG CTT C Gly Asn Lys Leu L 3215	
35			AGT CCT TTA TCA C Ser Pro Leu Ser L 3230	
40	AAA AGG AAG TC Lys Arg Lys Se	T GTT TCC ACA CCT r Val Ser Thr Pro 3240	GTC TCA GCC CAG A Val Ser Ala Gln M 3245	TG ACT TCA AAG 9981 et Thr Ser Lys 3250
45		y Glu Lys Glu Ile	GAT GAC CAA AAG A Asp Asp Gln Lys A 3260	
50			AGA CTG CCT TTA C Arg Leu Pro Leu P 32	
50			CCG GCT GCA CAG A Pro Ala Ala Gln L 3295	
55	Pro Pro Arg Se	er Cys Gly Thr Lys	TAC GAA ACA CCC A Tyr Glu Thr Pro I 3310	
	3300	3305		

5			Leu					Ile				GAA Glu	Leu				10269
		Thr					Ser					GAA Glu					10317
10	Ser					Thr					Thr	AGT Ser 3375				TAT Tyr	10365
15					Arg					Ser		ATC Ile			Gln		10413
20				Ala					Cys			AAT Asn		Gln			10461
			Thr	AAA Lys 3415				TAA									10485
25																	
		1 :		INE						NO:5	5 :						
30		()	(A) (B) (C)	LENC TYPE STRA	ETH: E: an ANDEI	3418 mino ONESS	acio acio S: si	ino a i ingle	cids	i							
35				OLEC RAGME			. •										
		(2	(i) §	BEQUE	ENCE	DESC	RIP	CION:) ID	NO:5	5:					
40	Met	-	•	_					SEÇ				Glu	Ile	Phe	Lys	
40	1	Pro	Ile	Gly Asn	Ser 5	Lys	Glu	Arg	SEC Pro Gly	Thr 10	Phe	Phe Ser		Asn	15		
	1 Thr	Pro Arg	Ile Cys	Gly Asn 20	Ser 5 Lys	Lys Ala	Glu Asp	Arg Leu	Pro Gly 25	Thr 10 Pro	Phe Ile	Phe	Leu Glu	Asn 30	15 Trp	Phe	
40 45	1 Thr Glu	Pro Arg Glu	Ile Cys Leu 35	Gly Asn 20 Ser	Ser 5 Lys Ser	Lys Ala Glu	Glu Asp Ala	Arg Leu Pro 40	Pro Gly 25 Pro	Thr 10 Pro Tyr	Phe Ile Asn	Phe Ser	Leu Glu 45	Asn 30 Pro	15 Trp Ala	Phe Glu	
	1 Thr Glu Glu Pro	Pro Arg Glu Ser 50	Ile Cys Leu 35 Glu	Gly Asn 20 Ser	Ser 5 Lys Ser Lys	Lys Ala Glu Asn Ser	Glu Asp Ala Asn 55	Arg Leu Pro 40 Asn	Pro Gly 25 Pro Tyr	Thr 10 Pro Tyr Glu	Phe Ile Asn Pro	Phe Ser Ser Asn	Leu Glu 45 Leu	Asn 30 Pro	15 Trp Ala Lys	Phe Glu Thr	
	1 Thr Glu Glu Pro 65	Pro Arg Glu Ser 50 Gln	Ile Cys Leu 35 Glu Arg	Gly Asn 20 Ser His	Ser 5 Lys Ser Lys Pro Gly	Lys Ala Glu Asn Ser 70	Glu Asp Ala Asn 55 Tyr	Arg Leu Pro 40 Asn Asn	Pro Gly 25 Pro Tyr Gl	Thr 10 Pro Tyr Glu Leu	Phe Ile Asn Pro Ala 75	Phe Ser Ser Asn	Leu Glu 45 Leu Thr	Asn 30 Pro Phe Pro	15 Trp Ala Lys Ile Val	Phe Glu Thr Ile	
45	Thr Glu Glu Pro 65 Phe	Pro Arg Glu Ser 50 Gln Lys	Ile Cys Leu 35 Glu Arg	Gly Asn 20 Ser His Lys Gln Lys	Ser 5 Lys Ser Lys Pro Gly 85	Lys Ala Glu Asn Ser 70 Leu	Glu Asp Ala Asn 55 Tyr	Arg Leu Pro 40 Asn Asn Leu	Pro Gly 25 Pro Tyr Gl Pro	Thr 10 Pro Tyr Glu Leu 90	Phe Ile Asn Pro Ala 75 Tyr	Phe Ser Ser Asn 60 Ser	Leu Glu 45 Leu Thr	Asn 30 Pro Phe Pro Pro	15 Trp Ala Lys Ile Val 95	Phe Glu Thr Ile 80 Lys	
4 5	Thr Glu Glu Pro 65 Phe Glu	Pro Arg Glu Ser 50 Gln Lys Leu	Ile Cys Leu 35 Glu Arg Glu Asp	Gly Asn 20 Ser His Lys Gln Lys 100	Ser 5 Lys Ser Lys ro Gly 85 Phe	Lys Ala Glu Asn Ser 70 Leu	Glu Asp Ala Asn 55 Tyr Thr Leu	Arg Leu Pro 40 Asn Asn Leu Asp	Pro Gly 25 Pro Tyr Gl Pro Leu 105	Thr 10 Pro Tyr Glu Leu Leu 90 Gly	Phe Ile Asn Pro Ala 75 Tyr	Phe Ser Ser Asn 60 Ser	Leu Glu 45 Leu Thr Ser Val	Asn 30 Pro Phe Pro Pro 110	15 Trp Ala Lys Ile Val 95 Asn	Phe Glu Thr Ile 80 Lys Ser	
45	Thr Glu Glu Pro 65 Phe Glu Arg	Pro Arg Glu Ser 50 Gln Lys Leu His	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115	Gly Asn 20 Ser His Lys Gln Lys 100 Ser	Ser 5 Lys Ser Lys Pro Gly 85 Phe Leu	Lys Ala Glu Asn Ser 70 Leu Lys Arg	Glu Asp Ala Asn 55 Tyr Thr Leu Thr	Arg Leu Pro 40 Asn Asn Leu Asp Val 120	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys	Thr 10 Pro Tyr Glu Leu 90 Gly	Phe Ile Asn Pro Ala 75 Tyr Arg Lys	Phe Ser Ser Asn 60 Ser Gln Asn Met	Leu Glu 45 Leu Thr Ser Val Asp 125	Asn 30 Pro Phe Pro Pro 110 Gln	15 Trp Ala Lys Ile Val 95 Asn	Phe Glu Thr Ile 80 Lys Ser Asp	
4 5	Thr Glu Glu Pro 65 Phe Glu Arg Asp	Pro Arg Glu Ser 50 Gln Lys Leu His Val	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115 Ser	Gly Asn 20 Ser His Lys Gln Lys 100 Ser Cys	Ser 5 Lys Ser Lys Pro Gly 85 Phe Leu	Lys Ala Glu Asn Ser 70 Leu Lys Arg Leu His	Glu Asp Ala Asn 55 Tyr Thr Leu Thr Leu 135	Arg Leu Pro 40 Asn Asn Leu Asp Val 120 Asn	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys Ser	Thr 10 Pro Tyr Glu Leu 90 Gly Thr	Phe Ile Asn Pro Ala 75 Tyr Arg Lys Leu Arg	Phe Ser Ser Asn 60 Ser Gln Asn Met	Leu Glu 45 Leu Thr Ser Val Asp 125 Glu	Asn 30 Pro Phe Pro Pro 110 Gln ser	15 Trp Ala Lys Ile Val 95 Asn Ala Pro	Phe Glu Thr Ile 80 Lys Ser Asp Val	
4 5	Thr Glu Glu Pro 65 Phe Glu Arg Asp Val 145	Pro Arg Glu Ser 50 Gln Lys Leu His Val 130 Leu	Ile Cys Leu 35 Glu Arg Glu Asp Lys 115 Ser Gln	Gly Asn 20 Ser His Lys Gln Lys 100 Ser Cys	Ser 5 Lys Ser Lys Fro Gly 85 Phe Leu Pro	Lys Ala Glu Asn Ser 70 Leu Lys Arg Leu His	Glu Asp Ala Asn 55 Tyr Thr Leu Thr Leu 135 Val	Arg Leu Pro 40 Asn Asn Leu Asp Val 120 Asn Thr	Pro Gly 25 Pro Tyr Gl Pro Leu 105 Lys Ser Pro	Thr 10 Pro Tyr Glu Leu 90 Gly Thr Cys	Phe Ile Asn Pro Ala 75 Tyr Arg Lys Leu Arg 155	Phe Ser Ser Asn 60 Ser Gln Asn Met Ser 140	Leu Glu 45 Leu Thr Ser Val Asp 125 Glu Lys	Asn 30 Pro Phe Pro Pro 110 Gln Ser	15 Trp Ala Lys Ile Val 95 Asn Ala Pro	Phe Glu Thr Ile 80 Lys Ser Asp Val Val 160	

				180					185					190		
	Ser	Trp	Ser 195	Ser	Ser	Leu	Ala	Thr 200	Pro	Pro	Thr	Leu	Ser 205	Ser	Thr	Val
5	Leu	Ile 210	Val	Arg	Asn	Glu	Glu 215	Ala	Ser	Glu	Thr	Val 220	Phe	Pro	His	Asp
	Thr 225	Thr	Ala	Asn	Val	Lys 230	Ser	Tyr	Phe	Ser	Asn 235	His	Asp	Glu	Ser	Leu 240
10	_			-	245					250					Asn 255	
				260					265					270	Gly	
			275					280					285		Met	
15		290					295					300			Ser	
	305	_				310					315				Asn	320
20					325					330					Glu 335	
				340					345					350	Lys	
25			355					360		_		_	365		Asp	
23		370				_	375					380			Thr	
	385					390			_		395				His	400
30		-			405					410					415 Thr	
			-	420					425					430	Pro	
35			435					440					445		Thr	
		450				_	455					460			Asp	
	465 Ile	Leu	Ala	Val	Lys	470 Gln	Ala	Ile	Ser	Gly	475 Thr	Ser	Pro	Val	Ala	480 Ser
40	Ser	Phe	Gln	Gly	485 Ile	Lys	Lys	Ser	Ile	490 Phe	Arg	Ile	Arg	Glu	495 Ser	Pro
	Lys	Glu		500 Phe	Asn	Ala	Ser		505 Ser	Gly	His	Met		510 Asp	Pro	Asn
45	Phe		515 Lys	Glu	Thr	Glu		520 Ser	Glu	Ser	Gly		525 Glu	Ile	His	Thr
	Val 545	530 Cys	Ser	Gln	Lys	Glu 550	535 Asp	Ser	Leu	Cys	Pro 555	540 Asn	Leu	Ile	Asp	Asn 560
50		Ser	Trp	Pro	Ala 565		Thr	Thr	Gln	Asn 570		Val	Ala	Leu	Lys 575	
	Ala	Gly	Leu	Ile 580		Thr	Leu	Lys	Lys 585		Thr	Asn	Lys	Phe 590	Ile	Tyr
	Ala	Ile	His 595		Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605	Pro	Lys	Asp
55	Gln	Lys 610	Ser	Glu	Leu	Ile	Asn 615	Сув	Ser	Ala	Gln	Phe 620	Glu	Ala	Asn	Ala
	625					630					635				Leu	640
60					645					650					Pro 655	
	Leu	Ser	Leu	Thr 660	Ser	Ser	Phe	Gly	Thr 665	Ile	Leu	Arg	Lys	Сув 670	Ser	Arg

			675					680					685			Tyr
5		690					695					700		Ile		
	705		_			710					715			Glu		720
		_		_	725					730				Leu	735	
10				740					745					750		
			755		_			760					765	Ala Leu		
15		770					775					780				Gly
	785		_	=	=	790					795					800
			_		805					810				Pro	815	
20	_			820					825					Asn 830		
	Leu	Leu	835	Pro	Glu	гЛа	ıyr	Met 840	Arg	vaı	Ala	ser	845	Ser	Arg	гуя
25		850					855					860		Lys		
	Glu 865	Glu	Thr	Thr	Ser	Ile 870	Ser	Lys	Ile	Thr	Val 875	Asn	Pro	Asp	Ser	Glu 880
					885					890				Val	895	
30	Glu	Arg	Asn	Asn 900	Leu	Ala	Leu	Gly	Asn 905	Thr	Lys	Glu	Leu	His 910	Glu	Thr
	_		915	•				920					925	Thr		
35		930					935					940		Ser		
	Lys 945	Asp	Leu	Val	Tyr	Val 950	Leu	Ala	Glu	Glu	Asn 955	Lys	Asn	Ser	Val	Lys 960
				_	965					970				Asp	975	
40				980					985					Met 990		
	_		995					1000)				1005			
45		1010)				1015	5				1020)	His		
	102	5		_		1030)				1035	5				Thr 104
				-	1045	5				1050)			Asp	1055	i
50	Lys	Lys	Leu	Ser 1060		Pro	Gln	Ser	Ile 1065		Thr	Val	Ser	Ala 1070		Leu
	Gln	Ser	Ser 1075		Val	Val	Ser	Asp 1080		Lys	Asn	Ser	His 1089	Ile 5	Thr	Pro
55	Gln	Met 1090		Phe	Ser	Lys	Gln 109		Phe	Asn	Ser	Asn 1100		Asn	Leu	Thr
	Pro 110		Gln	Lys	Ala	Glu 1110		Thr	Glu	Leu	Ser 1115		Ile	Leu	Glu	Glu 112
•	Ser	Gly	Ser	Gln	Phe 112		Phe	Thr	Gln	Phe 1130		Lys	Pro	Ser	Tyr 1139	
60	Leu	Gln	Lys	Ser 114	Thr		Glu	Val	Pro 114	Glu		Gln	Met	Thr 1150		Leu
	Lys	Thr	Thr			Glu	Сув	Arg			Asp	Leu	His	Val	Ile	Met

			1155					1160					1165	_,	a 1	a)
		4 4 7 4					1175					Lys 1180				
5	Thr	Val	Glu			1190	1				TTAD	Leu				120
	Asn	Lys			1200					1210)	Asn			1211	
10				1226	Ser	Ala			1225)		Asn		1230		
10			1225	Lys	Ala			1240)			Ile	1240)		
		1200	Thr	Ser			1255	;				Leu 1260	ı			
15	1000	His	qaA			1270)				1275					120
	Lys	Thr			1201	=				1290)	Leu			123.	,
20				1200	Thr	Thr			1305	•		Glu		131	,	
20			1211	Asn	Thr			1320)			Tyr	1323	,		
		1226	^				133!	5				Asp 1340	,			
25		_				125	Ω				135	Leu 5				130
	Gln	His			126	5				137	0	Phe			13/	,
30				138	Lys n	Glu			138	5		Thr		139	,	
30			120	=				140	٥			Ser	140	5		
		7 4 7	^				141	5				Asp 142	U			
35	440	_				343	ი				143					7.4.4
					144	5				145	0	Gln			143	2
40				116	Λ				146	5		Ser		14/	U	
			7.47					748	D			Asp	140	-		
		7 4 0	۸.				149	15				120	U			Val
45	160	\ =				151	.0				151	.5				Thr 152
					153	25				153	30				100	
50				15/	ı۸				154	15				723	U	Gly
			16	55				156	50				120	• •		Lys
		3 5 7	70				15	75				126	50			Glu
55	15	0.5				15	90				15	95				160
	As	p Ly			16	05				16	10				10.	
60				16	20				163	25				Το.	5 U	s Ser
	11	e Ph		u Ly 35	s Va	1 Ly	s Va	1 Hi 16	s Gl	u As	n Va	l Glı	16	s G11 45	a Th	r Ala

			_	Ala			7666					1000				
	Glu	Asn	Ser	Ala	Leu	Ala	Phe	Tyr	Thr	Ser	Cys	Ser	Arg	Lys	Thr	Ser 168
5		-		Thr		1670)				TO 12					
					3 6 0 0	=				1690	}				1033	
	Ile	Phe	Asp	Gly	Gln	Pro	Glu	Arg	Ile 1705	Asn	Thr	Ala	Asp	Tyr 1710	Val	Gly
10	Asn	Tyr	Leu 171	170 Tyr	Glu	Asn	Asn	Ser 1720	Asn	Ser	Thr	Ile	Ala 1725	Glu	Asn	Asp
			His	Leu			172	Gln	Asp			1/40				
	Met	Ser	Asn	Ser	Tyr	ser	Tyr	His	Ser	Asp	Glu	Val	Tyr	Asn	Asp	Ser
15	204	_				1751	n .				1/5	,				
				Ser	776	_				1773	0				 / ~	,
				Glu 178	Λ				1785	•				1/50	,	
20			170	a Asp 95				1800	0				TOOS	_		
	Cys	Va]	. Glu	ı Glu	Leu	Val	Thr	Ser	Ser	Ser	Pro	Cys 1820	Lys	Asn	гÀа	Asn
		181	LO	e Lys	Tou	Car	181	5 Ser	Agn	Ser	Asn			Glu	Val	Gly
25	100	_				183	n				183	>				TOZ
2.5	Pro	Pro	Ala	a Phe	Arg	Ile	Ala	Ser	Gly	Lys	Ile	Val	Cys	Val	Ser 185	His
					184	5				185	U				105.	,
				e Lys 186	٠.				186	5				TO 1	U	
30				s Glu	Asn			7 2 2	Lys ^	Ser			100.	J		
	Ile	Me	L Al	75 a Gly	, Сув	Tyr	Glu	Ala	Leu	Asp	Asp	Ser	Glu	Asp	Ile	Leu
		7.0	2				189	15				TOO	U			
			n Se	r Let	1 Asp	Asn 191		Glu	Cys	ser	191	лтэ 5	261	111.0	Lys	192
35	190)5 Δ1.	a As	p Ile	Gln	Ser	Glu	Glu	Ile	Leu	Gln	His	Asn	Gln	Asn	Met
					100					193	U				100	_
				u Glu 194	. ^				144	-					•	
40	Glu	ı Th		r Ası	, Ile	Cys	Lys	Cys 196	Ser	Ile	GIY	гÀв	196	. mis	гур	361
	Va.		r Se	55 r Ala	a Asr	1 Thr	Cys	Gly	Ile	Phe	ser	Thr 198	Ala	Ser	Gly	Lys
	Ç o	19 - Va	70 1 G1	n Va	l Sei	- Asr	Ala	a Ser	Leu	Glr	. Asn	Ala	Arg	Gln	Val	Phe 200
45	10	0.5				190	ลก				133	· 🗢				
	Se	r Gl	u Il	e Gl	u Asj	Ser	Thi	r Lys	Glr	val 201	l Phe	Ser	ГЛа	vai	201	Phe
	Ly	s Se	r As			05 ຣ ລະ:	c As	p Glr	1 Let 202	Thi	. Arg	g Glu	Glu	Asn 203	Thr	Ala
ΕO	T 3	~ N~	-~ T	20 r Pr	20 ດ ເ∃າ	u His	s Le	u Ile	Sei	c Glr	n Lys	Gly	Phe			Asn
50			2.0	125				204	10				204	13		
		20	\EA				20.	55				206				Lys
	Gl	n Va	al Se	er Il	e Le	u Gl	u Se	r Se	r Le	u Hi	s Lys	s Val	. Lys	s G1)	, val	Leu 208
55	20	65_				20 11	70 ○	α ጥከ [.]	r Gli	n Hi	201 s Set		Hi	s Ту	c Sei	Pro
					20	25				20	90				20:	,,
	Th	r Se	er A	rg Gl	n As	n Va	l Se	r Ly	s Il	e Le	u Pro	o Arg	y Vai	l As	Ly:	a Arg
				21	An .				21	05				21.	10	
60			2	115				21	20				21.	25		r Lys
	G]	u P	he L	ys Le	eu Se	r As	n As	n Le	u As	n Va	1 G1	u Gl	y G1	y Se	r Se	r Glu

		2130)				2135	;				2140				
	2145	Asn	His			2150)				2155)		Phe		210
5	qaA	Lys			2165					2170)			Val	Z 1 / 3)
				2180	١				2185	5				Val 2190	,	
10				-				2200)				220:	Lys		
		2216	^				2219	3				2220	,	Tyr		
	222	=				2230)				2235	•		Asp		224
15					2249					2250)			Phe	225	,
				2250	^				226	5				Gly 2270	,	
20			227	=				2280)				228	rys Pys		
		220	^				229	5				2300	,	Arg		
25	220	=				231	O.				231	5		Pro		232
25					232	5				233	0			Thr	233	9
				224	0				234	5			His	Leu	U	Leu
30			236	5			Leu	236 Ala	0			His	Pro	2		Gln
	Val	237 Ser	0 Ala	Thr	Arg	Asn	237 Glu	5 Lys	Met	Arg	His	238 Leu		Thr	Thr	Gly 240
35	238 Arg	5 Pro	Thr	. Lys	Val	239 Phe	o Val	Pro	Pro	Phe	239 Lys		Lys	Ser	His 241	Phe
	His	Arg	val	Glu 242	240 Gln	cys Cys	Val	Arg	Asn 242	Ile	Asn	Leu	Glu	Glu 243	Asn	Arg
40			242	Asn	-le			244	Gly 0	Ser			244	-		Lys
40		245	Asp	Asn			245	Gln	Phe			246	U			Gln
	240	Ala	a Ala			247	, O				247	5				Leu 248
45	Ile	e Thr			248	5				249	, O				247	
				250	0.0				250)5				251	U	Leu
50			251	15				252	20				254	:5		. Gly - Glv
		25	30				253	35				254	U			Gly
	25	1 E				255	50				255	25				Phe 256 Gly
55					250	55				251	70				25	75 n Asp
				25	80				25	85				253	, ,	r Pro
60			25	95				260	00				26	U5		s Tyr
	GI		10				26	15		_		262	20			

	Arg	Trp	Ile	Ile	Trp	Lys	Leu	Ala	Ala	Met	Glu	Сув	Ala	Phe	Pro	шув 264
						2620					2635					
5					2015					207U	Arg					
3				~~~	Asp	Thr			7007	1	Ser					
	-			Met	Glu			2680	1		Ala		2003			
10			Ser	qaA			750	-			Asn	2100	,			
		Asn	Lys								Lys 2715					
		Thr				Tyr	Ala				Gln					
15				204	^	Asn			J 14'	Thr	Val				•	
				Ala	Glu			276	()		qaA		2/0-	,		
20			Pro	Glu			277	E-			Ser	410	,			
	Pro	Ala	Arg	Trp	Tyr	Thr	Lys	Leu	Gly	Phe	Phe	Pro	Asp	Pro	Arg	Pro 280
		_				279	n				412	,				
	Phe	Pro	Leu	Pro	Leu	Ser	Ser	Leu	Phe	281	Asp	GIY	GIY	ASII	281	5
25	_		•	11-1	280	5 Tla	Gln	Δra	Δla	Tvr	Pro	Ile	Gln	Trp		
				000	^				282	5				205	•	
				Ser	Gly			784	()		Asn		207	,		
30			Ala	Ala			- 7 0 5				Gln	200	•			
	Leu	Phe	Thr	Lys	Ile	Gln	Glu	ı Glu	Phe	Glu	Glu	His -	GIu	GIU	Asn	288
		_				207	^				20/	3				
35											Thr					
											Ala					
											Glu					
40								2 1			Lys	227				
						295	50				277	3				Gln 296
45										74	/ LI				,	
				200	0.0				2.91	15				4	, .	Pro
			200	n <i>c</i>				30	00				300	, ,		Ile
50	Ту	r Hi	s Le	u Al	a Th	r Se	r Ly	s Se	r Ly	s Se	r Lys	Se:	c Gli	Arg	j Ala	Asn
			4 ^				าวก	15				30.	۷۷			val
	20	2 -				30	30				303	99				
	se	r As	p Gl	u Il			e Gl	n Il	е Ту	r Gl:	n Pro	Ar	g GI	1 Pro	30!	ı His
55		_	_	-1	30	45	~ D~	- Ac	n Dh	30 30 a	ou n Pro	s Se	r Cv	s Se		val
				2.0	<i>-</i> ^ _				4 ()	65					, .	
								3.0	H()				30	U J		u Ala
60		2.0	e Va	1 Ту			3.0	195				27	00			e Lys
	Ph	ie Tr	p Il	e As	sp Le	u As	n G	lu As	p Il	e Il	e Ly	s Pr	o Hi	s Me	t Le	u Ile

	222	
	3105 3110 3115 312	
	alla Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Giy Leu Leu	
5	The Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu	
ט		
	Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn	
	3155 Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu 3180	
	2175	
10	3170 His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser 320	
	3185 3190 3195 3195 3195 3195 3195	
	Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu 3205 3210 3215	
15	Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu	
	Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met	
	Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn	
20	3765	
20	Circ Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro	
	3270 34/3	
	Pro Pro Val Ser Pro Ile Cvs Thr Phe Val Ser Pro Ala Ala Gin Lys	
	205 3290	
٥.5	Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile	
25	2220	
	Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe	
	Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu	
30	3330 3335 3340 Ser Thr Gly Glu Lys	
	Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys 3350 3350 3350 3350	
	2260 2333 277	
	Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser	
	22CE 33/V	
35	Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys	
	2000	
	Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys	
	3395 3400 5403	
	Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile	
40	3410 3415	
	(2) INFORMATION FOR SEQ ID NO:6:	
	(i) SEQUENCE CHARACTERISTICS:	
45	(A) LENGTH: 10485 base pairs	
43	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGI: linear	
	(b) Topologi. These	
	(ii) MOLECULE TYPE: cDNA	
50		
	(ix) FEATURE:	
	Coding Company	
	(A) NAME/KEY: Coding Sequence	
	(B) LOCATION: 22910482	
55	(D) OTHER INFORMATION: BRCA2 (OMI2)	
	THE PROPERTY OF THE WOLF.	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
	THE PROPERTY OF THE PROPERTY O	6
	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC	12
60	The second concern the second concern the second control of the se	18
	THE PROPERTY OF COCCOCCC CONTINUES OF CONTINUES OF CONTINUES OF COCCOCCC	37
	ACAGATTTGT GACCGGCGCG GTTTTGTGT GGAATATCGT AGGTAAAA ATG CCT ATT 2 CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAA ATG CCT ATT 2	ر د.

Met Pro Ile

														1			
5	GGA Gly	TCC Ser 5	AAA Lys	GAG Glu	AGG Arg	CCA Pro	ACA Thr 10	TTT Phe	TTT Phe	GAA Glu	ATT Ile	TTT Phe 15	AAG Lys	ACA Thr	CGC Arg	TGC Cys	285
10	AAC Asn 20	AAA Lys	GCA Ala	GAT Asp	TTA Leu	GGA Gly 25	CCA Pro	ATA Ile	AGT Ser	CIT Leu	AAT Asn 30	TGG Trp	TTT Phe	GAA Glu	GAA Glu	CTT Leu 35	333
15	TCT Ser	TCA Ser	GAA Glu	GCT Ala	CCA Pro 40	CCC Pro	TAT Tyr	AAT Asn	TCT Ser	GAA Glu 45	CCT Pro	GCA Ala	GAA Glu	GAA Glu	TCT Ser 50	GAA Glu	381
2.0	CAT His	AAA Lys	AAC Asn	AAC Asn 55	AAT Asn	TAC Tyr	GAA Glu	CCA Pro	AAC Asn 60	CTA Leu	TTT Phe	AAA Lys	ACT Thr	CCA Pro 65	CAA Gln	AGG Arg	429
20	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	AAT Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
25	C AA Gln	GGG Gly 85	CTG Leu	ACT Thr	CTG Leu	CCG Pro	CTG Leu 90	TAC Tyr	CAA Gln	TCT Ser	CCT Pro	GTA Val 95	AAA Lys	GAA Glu	TTA Leu	Asp	525
30	AAA Lys 100	TTC Phe	AAA Lys	TTA Leu	GAC Asp	TTA Leu 105	GGA Gly	AGG Arg	AAT Asn	GTT Val	CCC Pro 110	AAT Asn	AGT Ser	AGA Arg	CAT His	AAA Lys 115	573
35	AGT Ser	CTT Leu	CGC Arg	ACA Thr	GTG Val 120	AAA Lys	ACT Thr	AAA Lys	ATG Met	GAT Asp 125	CAA Gln	GCA Ala	GAT Asp	GAT Asp	GTT Val 130	TCC Ser	621
40	TGT Cys	CCA Pro	CTT Leu	CTA Leu 135	AAT Asn	TCT Ser	TGT Cys	CTT Leu	AGT Ser 140	Glu	AGT Ser	CCT Pro	GTT Val	GTT Val 145	CTA Leu	CAA Gln	669
40 ·	TGT Cys	ACA Thr	CAT His	GTA Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	TGT Cys	GGG	AGT Ser	717
45	TTG Leu	TTT Phe	His	ACA Thr	CCA	AAG Lys	TTT Phe 170	Val	Lys	GGT	CGT Arg	CAG Gln 175	Thr	CCA Pro	AAA Lys	CAT His	765
50	11e	e Sei	r Glu	ı Ser	Leu	185	Ala	Glu	ı Val	. Asp	190) Asp	Met	ser	Trp	Ser 195	813
55	Sei	c Se	r Lei	ı Ala	200	Pro	Pro	Thi	. Leu	205	Ser	Thr	· Val	Leu	210		861
60	Arg	aA g	n Gl	1 Glu 219	a Ala	a Sei	c Glu	ı Thi	r Val 220	l Phe	e Pro) His	a Asp	225	Thi	GCT Ala	909
00	AA As:	T GT n Va	G AA	A AGO	TA'	r Tr	TCC Se:	C AA'	r CA	r GA' s Asj	r GAZ p Gli	A AG'	r CTC	AAC 1 Lys	AAI Lys	TAA A	957

230 235 240

			230					:	235					2	240				
5	GAT Asp	AGA Arg 245	TTT Phe	ATC Ile	GCT Ala	T TO	er \	GTG A Val ' 250	ACA Thr	GAC Asp	AGT Ser	GAA Glu	· A	AC 1 sn 5	ACA I	AAT Asn	CAA Gln	AGA Arg	1005
10	GAA Glu 260	GCT Ala	GCA Ala	AGT Ser	CA'	s G	GA ' ly 1 65	TTT Phe	GGA Gly	AAA Lys	ACA Thr	TCA Ser 270		GG I	AAT Asn	TCA Ser	TTT Phe	AAA Lys 275	1053
	GTA Val	AAT Asn	AGC Ser	TGC Cys	AA Ly 28	s A	AC sp	CAC His	ATT Ile	GGA Gly	AAG Lys 285	TC# Ser	A A	TG let	CCA Pro	AAT Asn	GTC Val 290	CTA Leu	1101
15	GAA Glu	GAT Asp	GAA	GT Va: 29	1 Ту	T G r G	AA lu	ACA Thr	GTT Val	GTA Val 300	GAT Asp	AC(C T	CT	GAA Glu	GAA Glu 305	GAT Asp	AGT Ser	1149
20	TTT Phe	TC/ Ser	TTA Lev	з Су	r TT s Ph	т 1 е S	CT Ser	AAA Lys	ТСТ Сув 315	AGA Arg	ACA Thr	AA Ly	A A	TAJ ne	CTA Leu 320	CAA Gln	AAA Lys	GTA Val	1197
25	AGA Arg	ACT Thi	s Se	C AA	G AC s Th	T A	AGG Arg	AAA Lys 330	TA2 TY2	ATT Ile	TTC Phe	CA'	9 (GAA Glu B35	GCA Ala	AAC Asn	GCT Ala	GAT Asp	1245
30	GAA Glu 340	Су	r GA s Gl	A AA u Ly	A TO	r I	AAA Lys 345	AAC Asn	CAA Gln	GTG Val	AA! Lys	GA Gl 35	u i	AAA Lys	TAC Tyr	TCA Ser	TTT	GTA Val 355	1293
	TCI Ser	GA G1	A GT u Va	G GA 1 Gl	u Pi	CA 2	AAT Asn	GAT Asp	ACT Thr	GAT Asp	CC2 Pro 369) re	A C	GAT Asp	TCA Ser	AAT Asn	GTA Val 370	GCA Ala	1341
35	CAT His	CA Gl	G AA n Ly	G CC s Pr	o Pl	rr (GAG Glu	AGT Ser	GGA Gly	AGT Ser	AS	C AA o Ly	A I	ATC Ile	TCC Ser	AAG Lys 385		GTT Val	1389
40	GT/ Val	A CC l Pr	G TC	r Le	rg G eu A	CC la	TGT Cys	GAA Glu	TGG Trp 395	se:	r CA.	A CI	A .	ACC Thr	CTT Leu 400	001	GG?	CTA Leu	1437
45	AA' As:	T G0 n G1 40	y A.	C C	AG A ln M	TG et	GAG Glu	AAA Lys 410	Ile	CC	C CT	A Ti	±u	CAT His 415	110	TC1	TC Se:	A TGI c Cys	1485
50	GA As 42	p G	AA Ai In A	AT A' sn I	rr T le S	CA	GAA Glu 425	Lys	A GAC	C CT	A TI u Le	u A	AC sp 30	ACA Thr	GAG Glu	AA(C AA n Ly	A AGA 8 Arg 435	•
	AA Ly	G A	AA G ys A	AT T sp P	he I	TT eu	ACT Thr	TCI Sei	A GAG	G AA u As	T TC n Se 44	E L	TG eu	CCA	CGT Arg	r AT	T TC e Se 45	T AGO r Sei	1581
55	CT Le	A C	CA A ro L	ys S	CA (er (GAG Glu	AA(G CC	A TT.	A AA u As 46	in G.	AG G Lu G	AA lu	AC! Thi	A GTO	G GT l Va 46		T AA	G 1629 s
60	A(A)	BA G	ge.	AA (lu (BAG (CAG Gln	CA' Hi	T CT s Le	T GA u Gl 47	u se	CT C.	AT A	CA hr	GA(C TG P Cy 48		T Cl	T GC au Al	A 1677 a

5	GTA Val	AAG Lys 485	CAG Gln	GCA Ala	ATA Ile	TCT Ser	GGA Gly 490	ACT Thr	TCT Ser	CCA Pro	GTG Val	GCT Ala 495	TCT Ser	TCA Ser	TTT Phe	CAG Gln	1725
-	GGT Gly 500		AAA Lys	AAG Lys	TCT Ser	ATA Ile 505	TTC Phe	AGA Arg	ATA Ile	AGA Arg	GAA Glu 510	TCA Ser	CCT Pro	TA8 YYY	GAG Glu	ACT Thr 515	1773
10	TTC Phe	AAT Asn	GCA Ala	AGT Ser	TTT Phe 520	TCA Ser	GGT Gly	CAT His	ATG Met	ACT Thr 525	GAT Asp	CCA Pro	AAC Asn	TTT Phe	AAA Lys 530	FÀ2 YYY	1821
15	Glu	Thr	Glu	GCC Ala 535	Ser	Glu	Ser	Gly	ьеи 540	GIU	TIE	uip	IIII	545	Cyb	202	1869
20	Gln	ГÀв	Glu 550		Ser	Leu	Cys	Pro 555	Asn	ьeu	116	Kap	560	O17	501		1917
25	Pro	Ala 565	Thr	ACC Thr	Thr	Gln	Asn 570	Ser	vai	Ala	Leu	575	Abii	AIG	UL,	200	1965
30	Ile 580	Ser	Thr	TTG Leu	Lys	Lys 585	Lys	Thr	Asn	ьys	590	116	1y1	AIG	110	595	2013
30	Asp	Glu	Thi	Ser	Tyr 600	Lys	Gly	Lys	Lys	605	PIO	пуъ	vaħ	0111	610		2061
35	Glu	ı Lev	ı Ile	Asn 615	Cys	Ser	Ala	Gin	620	GIU	Ата	Abii	AIG	625	-		2109
40	Pro	Let	63	r Phe	Ala	Asn	Ala	Asp 635	ser	GIŸ	Per	l Deu	640		-	GTG Val	2205
45	Ly	64	g Se	r Cya	s Ser	Glr	650	n Asp)	ser	GIU	l GI	655	2 1111	. Dec	. 501	TTA Leu	2253
50	Th:	r Se 0	r Se	r Phe	e Gly	7 Thi 665	r Ile	e Lei	ı Arg	d rAs	679	9 561	. ALC	y No.	, 01,	A ACA Thr 675	2301
	СУ	s Se	r As	n As:	n Th:	r Val 0	1 [1]	e Sei	r Gi	n AS] 68	5 ње 5	ц жы	, IA	Luy	69		2349
55	Lу	s Cy	s As	n Ly 69	s Gl 5	u Ly	s Le	u Gl	n Le [*]	u Pn	e II	e in	L PL	70	5	T GAT a Asp A AGC	2397
60	TC Se	T CT er Le	eu Se	A TG er Cy 10	C CT	G CA u Gl	G GA n Gl	A GG u Gl 71	y Gl	с TG n Су	r GA s Gl	и AS	n As 72	b br	o Ly	A AGC s Ser	2007

	WO	77/07	104														
	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	Ile	AAA Lys 730	GAA (Glu (GAG (GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
5	CCA Pro 740	GTA Val	CAA Gln	CAT His	TCA Ser	AAA Lys 745	GTG Val	GAA ' Glu '	TAC Tyr	Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
10	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
15	ACT Thr	CCT	ACT Thr	TCC Ser 775	AAG Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	AAC Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	TCT Ser	AGA Arg	2589
20	Gly	Lys	Glu 790		Tyr	Lys	Met	5er 795	qaA	гув	Leu	гув	800	ASII	AD.	- 7 -	2637
٥٣	GAA Glu	TCT Ser 805	Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	AAA Lys	TAA Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
25	GAT Asp 820	val	TGI Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
30	CCT Pro	GAZ	A AAA 1 Lys	TAC Tyr	ATG Met 840	AGA Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	TCA Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	FIIC	2781
35	AA(Ası	C CA	A AAC n Ası	ACA Thr	Asn	CTA	AGA Arg	GTA Val	ATC Ile 860	Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GIU	ACT	2829
40	Th	T TC.	A ATT	e Ser	AAA Lys	ATA Ile	ACT Thr	GTC Val 875	AAT Asn	CCA Pro	GAC Asp	: TCT	GAA Glu 880	GIU	CTT Leu	TTC Phe	2877
	Se	A GA r As 88	p Ası	r GAG	AAT Asn	AAT Asn	TTT Phe 890	. Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	ASI	GAA Glu	AGG Arg	AAT Asn	2925
45		n Le	T GC	T TT! a Le	A GGA 1 Gly	AAT AST 905	Thr	AAG Lys	GAA Glu	CTI Leu	CAT His 910	s GIV	ACA Thr	GAC Asp	TTC Lev	ACT Thr 915	2973
50) ТС Су	T GI s Va	AA AA 1 As	C GAA	A CCC u Pro 920	o Ile	TTC Phe	AAG Lys	AAC ASI	TC1 1 Se1 925	r Th	C ATO	GTT Val	TTA Let	TAT 1 Ty: 930	r GGA r Gly	3021
5	GA 5 As	C AC	CA GG ir Gl	T GA y As	р Гу	A CA	A GCA	A ACC	C CAI	n Va	G TC	A AT	r AAl e Ly:	A AAI s Lys 94!	S AB	r TTG p Leu	3069
6	Va	TT T	yr Va	TT CT al Le	T GC	A GA a Gl	G GA	G AAG u Ası 95!	n Ly	A AA' s As	T AG n Se	T GT. r Va	A AA 1 Ly: 96	B G1.	G CA n Hi	T ATA s Ile	3117
	A	AA A	TG A	CT CT	A GG	T CA	A GA	T TT	A AA	A TC	G GA	C AT	C TC	C TT	g aa	ATA T	3165

	Lys Met 965		Leu	Gly	Gln	Asp 970	Leu	Lys	Ser	Asp	Ile 975	Ser	Leu	Asn	Ile	
5	GAT AAA Asp Lys 980	ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	AAT Asn	AAT Asn	GAT Asp	TAC Tyr	ATG Met 990	AAC Asn	AAA Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
10	CTC TTA Leu Leu	GGT Gly	Pro	ATT Ile	TCA Ser	AAT Asn	CAC His	Ser	TTT Phe 1005	GGA Gly	GGT Gly	AGC Ser	Phe	AGA Arg 1010	ACA Thr	3261
15	GCT TCA Ala Ser	Asn	AAG Lys 1015	GAA Glu	ATC Ile	AAG Lys	Leu	TCT Ser L020	GAA Glu	CAT His	AAC Asn	Ile	AAG Lys 1025	AAG Lys	AGC Ser	3309
20	AAA ATG Lys Met	TTC Phe 1030	TTC Phe	AAA Lys	GAT Asp	Ile	GAA Glu 1035	GAA Glu	CAA Gln	TAT Tyr	Pro	ACT Thr L040	AGT Ser	TTA Leu	GCT Ala	3357
20	TGT GTT Cys Val 1045	Glu	ATT Ile	GTA Val	Asn	ACC Thr L050	TTG Leu	GCA Ala	TTA Leu	qaA	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
25	AGC AAG Ser Lys 1060	CCT Pro	CAG Gln	Ser	ATT Ile 1065	AAT Asn	ACT Thr	GTA Val	Ser	GCA Ala 1070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
30	GTA GTT Val Val	GTT Val	Ser	GAT Asp L080	TGT Cys	AAA Lys	AAT Asn	Ser	CAT His 1085	ATA Ile	ACC Thr	CCT Pro	Gln	ATG Met 1090	TTA Leu	3501
35	TTT TCC Phe Ser	Lys	CAG Gln 1095	GAT Asp	TTT Phe	AAT Asn	Ser	AAC Asn L100	CAT His	AAT Asn	TTA Leu	Thr	CCT Pro	AGC Ser	CAA Gln	3549
	AAG GCA Lys Ala	GAA Glu 1110	ATT Ile	ACA Thr	GAA Glu	Leu	TCT Ser 1115	ACT Thr	ATA Ile	TTA Leu	Glu	GAA Glu L120	TCA Ser	GGA Gly	AGT Ser	3597
40	CAG TTT Gln Phe 1125	Glu	TTT Phe	ACT Thr	Gln	TTT Phe 1130	AGA Arg	AAR Xaa	CCA Pro	Ser	TAC Tyr 1135	ATA Ile	TTG Leu	CAG Gln	AAG Lys	3645
45	AGT ACA Ser Thr 1140	TTT Phe	GAA Glu	Val	CCT Pro	GAA Glu	AAC Asn	CAG Gln	Met	ACT Thr 1150	ATC Ile	TTA Leu	AAG Lys	Thr	ACT Thr 155	3693
50	TCT GAG Ser Glu	GAA Glu	Cys	AGA Arg 1160	GAT Asp	GCT Ala	GAT Asp	Leu	CAT His 1165	GTC Val	ATA Ile	ATG Met	Asn	GCC Ala 1170	CCA Pro	3741
55	TCG ATT Ser Ile	Gly	CAG Gln 1175	GTA Val	GAC Asp	AGC Ser	Ser	AAG Lys 1180	CAA Gln	TTT Phe	GAA Glu	Gly	ACA Thr 1185	GTT Val	GAA Glu	3789
	ATT AAA	CGG Arg 1190	Lys	TTT Phe	GCT Ala	Gly	CTG Leu 1195	TTG Leu	AAA Lys	AAT Asn	Asp	TGT Cys 1200	AAC Asn	AAA Lys	AGT Ser	3837
60	GCT TCT Ala Ser	GGT Gly	TAT Tyr	TTA Leu	ACA Thr	GAT Asp	GAA Glu	AAT Asn	GAA Glu	GTG Val	GGG Gly	TTT Phe	AGG Arg	GGC Gly	TTT Phe	3885

1205 1210 1215

5				GTT TCT ACT GAA Val Ser Thr Glu 1230	
10			Ser Asp Ile	GNG AAT ATT AGT Glu Asn Ile Ser 245	
15				TCT TCA AGT AAA Ser Ser Ser Lys	
10		l Ser Met Phe		AAT CAT AAT GAT Asn His Asn Asp 1280	
20		s Asn Asn Lys		ATA TTA CAA AAT Ile Leu Gln Asn 1295	
25				ATT ACT GAA AAT Ile Thr Glu Asn 1310	
30			Asn Lys Tyr	ACT GCT GCC AGT Thr Ala Ala Ser 325	
35				TCA AGT AAA AAT Ser Ser Lys Asn	
		Lys Asp Glu		CTA TTT ACT GAT Leu Phe Thr Asp 1360	
40		Lys Leu Ser		ATG AAG GAG GGA Met Lys Glu Gly 1375	
45	Ile Lys Glu		Asp Leu Thr I	TTT TTG GAA GTT Phe Leu Glu Val 1390	
50			Asn Thr Ser A	AAT AAA GAA CAG Asn Lys Glu Gln 105	
55				TTT GAG ACT TCT Phe Glu Thr Ser 1	
33		Ala Ser Gly		AGT GTC GCC AAA Ser Val Ala Lys 1440	
60		Val Asn Phe		AAA CCA GAA GAA Lys Pro Glu Glu 1455	

5	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475	4653
	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
10	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
15	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
20	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA AFT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
25	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1555 1550	4893
	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570	4941
30	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
35	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
40	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
45	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	5133
F.0	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val Mis Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
50	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
55	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
60	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325

5	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
5	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
10	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
15	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
20	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
25	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
30	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
35	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
40	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
45	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
50	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
55	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
60	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA	6093

	Glu Lys 1940	Val Ser I	Lys Ile Ser 1945	Pro Cys	Asp Val 1950	Ser Leu Gl	u Thr Ser 1955	
5	GAT ATA Asp Ile	Cys Lys C	rGT AGT ATA Cys Ser Ile 960	GJA FAa	CTT CAT Leu His 1965	AAG TCA GT Lys Ser Va	C TCA TCT l Ser Ser 1970	6141
10	GCA AAT Ala Asn	ACT TGT G Thr Cys G 1975	GGG ATT TTT Gly Ile Phe	AGC ACA Ser Thr 1980	GCA AGT Ala Ser	GGA AAA TC Gly Lys Se 198	r Val Gln	6189
15	Val Ser	GAT GCT T Asp Ala S	CCA TTA CAP Ser Leu Gli	A AAC GCA Asn Ala 1995	AGA CAA Arg Gln	GTG TTT TC Val Phe Se 2000	T GAA ATA r Glu Ile	6237
20	GAA GAT Glu Asp 2005	AGT ACC A Ser Thr I	AAG CAA GTO Lys Gln Val 2010	Phe Ser	Lys Val	TTG TTT AA Leu Phe Ly 2015	A AGT AAC s Ser Asn	6285
20	GAA CAT Glu His 2020	TCA GAC O	CAG CTC ACA Gln Leu Thi 2025	A AGA GAA Arg Glu	GAA AAT Glu Asn 2030	ACT GCT AT Thr Ala Il	A CGT ACT e Arg Thr 2035	6333
25	CCA GAA Pro Glu	His Leu 1	ATA TCC CAN lle Ser Gli 040	Lys Gly	TTT TCA Phe Ser 2045	TAT AAT GT Tyr Asn Va	G GTA AAT l Val Asn 2050	6381
30	TCA TCT Ser Ser	GCT TTC T Ala Phe S 2055	TCT GGA TT Ser Gly Phe	AGT ACA Ser Thr 2060	GCA AGT Ala Ser	GGA AAG CA Gly Lys Gl 206	n Val Ser	6429
35	Ile Leu	GAA AGT T Glu Ser S	rcc TTA CAG Ser Leu His	C AAA GTT Lys Val 2075	AAG GGA Lys Gly	GTG TTA GA Val Leu Gl 2080	G GAA TTT u Glu Phe	6477
4.0	GAT TTA Asp Leu 2085	ATC AGA A	ACT GAG CAT Thr Glu His 2090	s Ser Leu.	His Tyr	TCA CCT AC Ser Pro Th	G TCT AGA r Ser Arg	6525
40	CAA AAT Gln Asn 2100	GTA TCA A	AAA ATA CT Lys Ile Leu 2105	r CCT CGT 1 Pro Arg	GTT GAT Val Asp 2110	AAG AGA AA Lys Arg As	C CCA GAG n Pro Glu 2115	6573
45	CAC TGT His Cys	Val Asn S	TCA GAA ATG Ser Glu Meg 120	Glu Lys	ACC TGC Thr Cys 2125	AGT AAA GA Ser Lys Gl	A TTT AAA u Phe Lys 2130	6621
50	TTA TCA Leu Ser	AAT AAC 1 Asn Asn 1 2135	TTA AAT GT Leu Asn Va	r GAA GGT l Glu Gly 2140	GGT TCT Gly Ser	TCA GAA AA Ser Glu As 214	n Asn His	6669
55	Ser Ile	AAA GTT 1 Lys Val 2 2150	TCT CCA TA	r CTC TCT r Leu Ser 2155	CAA TTT Gln Phe	CAA CAA GA Gln Gln As 2160	C AAA CAA p Lys Gln	6717
60	CAG TTG Gln Leu 2165	GTA TTA (GGA ACC AA Gly Thr Ly 217	s Val Ser	Leu Val	GAG AAC AT Glu Asn Il 2175	T CAT GTT e His Val	6765
00	TTG GGA Leu Gly	AAA GAA Lys Glu	CAG GCT TC Gln Ala Se	A CCT AAA r Pro Lys	AAC GTA Asn Val	AAA ATG GA Lys Met Gl	A ATT GGT u Ile Gly	6813

	2180				:	2185				:	2190					2195	
5				Thr		TCT Ser			Pro					Ile			6861
10			Thr			AAA Lys		Ser					Glu			GCA Ala	6909
15		Glu				GCT Ala	Phe					Glu					6957
13	Lys					GCC Ala					Phe						7005
20					Leu	TCA Ser 2265				Ile					Gly		7053
25	CCC Pro			Leu		GGA Gly			Ser					Leu			7101
30	GAA Glu		Asp			ATA Ile		Asn					Leu				7149
35	AAA Lys	Ser				GGC Gly	Thr					Arg					7197
55	His										Pro						7245
40	GAA Glu 2340				Ile					Phe					Gln		7293
45				Lys		CAT His			Glu					Glu			7341
50	TCA Ser		Asn			GTT Val		Gly					Gln				7389
55	ACA Thr	Arg				ATG _. Met	Arg					Thr					7437
<i>J J</i>	AAA Lys 2					Pro					Ser						7485
60	GAA Glu 2420				Arg					Glu					Lys		7533

5	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
10	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
20	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
15	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
20	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515	7773
25	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
30	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545	7869
30	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560	7917
35	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
40	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595	8013
45	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610	8061
50	CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625	8109
30	ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640	8157
55	AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655	8205
60	TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675	8253

5	ATG GAA AGG GA: Met Glu Arg Asp	GAC ACA C Asp Thr A 2680	Ala Ala Lys	ACA CTT GTT Thr Leu Val	CTC TGT GTT Leu Cys Val 2690	TCT 8301 Ser
J	GAC ATA ATT TCA Asp Ile Ile Ser 269	r Leu Ser I				
10	ACT AGT AGT GCA Thr Ser Ser Ala 2710			Ala Ile Ile		
15	GGG TGG TAT GCT Gly Trp Tyr Ala 2725	a Val Lys <i>l</i> 27	Ala Gln Leu 730	Asp Pro Pro 2735	Leu Leu Ala	Val
20	TTA AAG AAT GGG Leu Lys Asn Gly 2740	y Arg Leu 1 2745	Thr Val Gly	Gln Lys Ile 2750	Ile Leu His	Gly 2755
25	GCA GAA CTG GTG Ala Glu Leu Va		Pro Asp Ala			
25	GAA TCT CTT ATO Glu Ser Leu Met 2779	Leu Lys I				
30	TGG TAT ACC AAI Trp Tyr Thr Lys 2790			Asp Pro Arg		
35	CCC TTA TCA TCC Pro Leu Ser Ser 2805	Leu Phe S				
40	GTA ATT ATT CAM Val Ile Ile Gli 2820				Glu Lys Thr	
45	TCT GGA TTA TAG Ser Gly Leu Ty	r Ile Phe A	Arg Asn Glu		Glu Lys Glu	
	GCA AAA TAT GTO Ala Lys Tyr Va: 285	l Glu Ala (
50	AAA ATT CAG GAG Lys Ile Gln Glu 2870			Glu Glu Asn		
55	TAT TTA CCA TC Tyr Leu Pro Se 2885	r Arg Ala I				
60	GAT GGT GCA GA Asp Gly Ala Gl 2900				Ala Asp Pro	
	TAC CTT GAG GG	T TAT TTC	AGT GAA GAG	CAG TTA AGA	GCC TTG AAT	AAT 9021

	Tyr Leu Glu	ı Gly Tyr Phe 2920		Gln Leu Arg 2925	Ala Leu Asn Asn 2930	
5					ATC CAG TTG GAA Ile Gln Leu Glu 2945	
10		a Ala Met Glu		Gln Lys Glu	CAA GGT TTA TCA Gln Gly Leu Ser 960	
15		l Thr Thr Val			AGC TAT TCA AAA Ser Tyr Ser Lys	
20					CCA TCA TCA GAT Pro Ser Ser Asp 2995	
20			Glu Gly Lys		ATT TAT CAT CTT Ile Tyr His Leu 3010	9261
25					AAC ATA CAG TTA Asn Ile Gln Leu 3025	
30		Lys Lys Thr		Gln Leu Pro	GTT TCA GAT GAA Val Ser Asp Glu 040	9357
35		e Gln Ile Tyr			CAC TTC AGC AAA His Phe Ser Lys	9405
40					GTG GAC CTA ATA Val Asp Leu Ile 3075	9453
40			Val Lys Lys		GCC CCT TTC GTC Ala Pro Phe Val 3090	9501
45					AAG TTT TGG ATA Lys Phe Trp Ile 3105	9549
50		Glu Asp Ile		His Met Leu	ATT GCT GCA AGC Ile Ala Ala Ser 120	9597
55		Trp Arg Pro			CTT ACT TTA TTT Leu Thr Leu Phe	9645
60			Phe Ser Ala		GAG GGC CAC TTT Glu Gly His Phe 3155	9693
••					AAT ATT GAC ATA Asn Ile Asp Ile	9741

	WO 99/09164 FC1/0598/10	1703
	3160 3165 3170	
5	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
10	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
15	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
15	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
20	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	9981
25	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265	0029
30	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT 1 Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	0077
35	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	0125
33	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys 3300 3305 3310 3315	0173
40	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 1 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	0221
45	TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 1 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	0269
50	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 1 Asn Thr Gln Ala Leu Leu Ser Gly Ser inr Gly Glu Lys Gln Phe Ile 3350 3355 3360	0317
55	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 1 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	0365
	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG 1 Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3380 3385 3390 3395	0413

10461

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410

60

ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3418 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala

				240					345		Gln			330		
_	Ser	Phe		Ser	Glu	Val	Glu	Pro 360	Asn	Asp	Thr	qaA	Pro 365	Leu	Asp	Ser
5		270		His			375	Phe				380				
		Glu	Val			200					Trp 395					400
10					405	Ala				4 I U	Ile				110	
				420	Gln				425		Asp			430		
15			425	:				440			Glu		445			
13		450	Sei	Leu			455				Leu	460				
	400	Asr	Ly			470					Glu 475					100
20	Ile	Let			495					490	Thr				493	
				EAC	١				505		Arg			310		
25			E 1	=				520			His		525			
							535				Gly	540				
	C 4 C					รรถ					Pro 555					
30					565					570	Ser				313	
				E 0 (٦.				585		Thr			330		
35				_				600			Lys		605			
		~ 1	^				คาร				Gln	020				
		_				630	}				Asp 635 Asp					0.40
40					645	5				650	}				033	Arg
				66	Λ				665	5				0/0		Tyr
45								680	1				002	•		Pro
							69	•				,,,,,				Asp 720
5 0	70	_				710	3				/ 13	•				720 Ala
50					72	5				73	0				, 55	asp
				74	٥				74	5				/50	,	Thr
55			7	55				760	0				70:	•		L Met
		7.	7.0				77	5				781	J			s Gly
60	70					79	0				79	•				800 t Glu
60					20	۱5				81	.0				01	5 l Glu
	רי.	и В	טוו ט	TH M	-P ve	cy						•	_			

				820					825					830		
			Pro	Pro												
5			Phe				Thr	Asn								
		Glu	Thr			270	Ser				0/3					
1.0	Glu		Phe			Asn										
10			Asn	~~~	Leu				405							
			Thr 915	Сув				920					743			
15			Gly				0.2 F					240				
		Asp	Leu													
20	Gln	His	Ile		065					9/6						
			Ile	000					985					200		
			Gly 995					1 11111	11				100	_		
25			995 Thr 0													
		_				בחד	n				103	_				Thr 104
30			Ala		4 0 4	_				רנו ו	•				~~~	_
30					Lys	Pro										Leu
				Val	Val			108	()				100	~		Pro
35							100	15				110	•			Thr
							^				111					Glu 112
	110 Sei	c Gly	/ Ser	Glr	Phe	Glu	Phe	Thr	Glr	Phe 113	Arg	xaa	Pro	Ser	Tyr 113	lle 5
40	Leı	ı Glı	n Lys			Phe	e Glu	ı Val	. Pro	Gli	ı Asr	Gln	Met	Thr 115	Tl∈ 0	Leu
					c Glu			116	Asp	Ala			T T (, _		Met
45			~ ~	Se			77'	n Val	L Asy			776	, ,			Gly
			l Glu			11	g Ly:	s Phe			11:	73				120
50	As	n Ly			12	r Gl	у Ту:			12	ΤO					
50					r Se	r Al			12	25					,,,	r Glu
				n Ly	s Al			12	411				14	T		e Ser
55			u Th	r Se			17	55				14	00			r Lys
		s Hi	s As			17	70				12	/ 🕽				n Asp 128
60	Ly	s Th				0.5				12	: 9 0					n Asn 95
	As	sn I	le Gl		t Th	ır Th	ır Gl	y Th	r Ph	e Va 105	ıı Gl	u Gl	u 11	13	10	u Asn

											T	TT	mb	פות	λla	Ser
	Tyr	Lys	Arg	Asn	Thr	Glu	Asn	GIU	Asp	Asn	гав	LAL	1325	W10	ALU	001
			1315	5				1320)		_				T	Non.
	Arg	Asn	Ser	His	Asn	Leu	Glu	Phe	qaA	Gly	Ser	Asp	ser	ser	гур	ABII
5			•				1225					1341	,			
_	Asp	Thr	Val	Cvs	Ile	His	Lys	Asp	Glu	Thr	Asp	Leu	Leu	Phe	Thr	Asp
						1251	١.				135:	•				100
	7345	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	Asn	Tla	Chra	T.e.11	Lvs	Leu	Ser	Glv	Gln	Phe	Met	Lys	Glu	Gly
	GIII	UIP	ASII	110	1365	= =====================================	_,_			137	D				1375	;
			Gln		1303	, ,		T 011	Car	Aen	T.e.11	Thr	Phe	Leu	Glu	Val
10	Asn	Thr	Gln			GIU	Asp	Lea	1385	- Tob	ДСИ			1390	1	
				1380)	_					mb	C	N c m			Gln
	Ala	Lys	Ala	Gln	Glu	Ala	Cys	His	GIY	Asn	THE	261	A511	·	014	0111
			1399	5				1400)		_	_	1405		mh	Cox
	Leu	Thr	Ala	Thr	Lys	Thr	Glu	Gln	Asn	Ile	ГÀг	Asp	Pne	GIU	THE	Set
15		944	^				7479	•				1421	,			
	Asp	Thr	Phe	Phe	Gln	Thr	Ala	Ser	Gly	Lys	Asn	Ile	Ser	Val	Ala	гуs
		-				7/2/	`				143	5				7.4.4
	C111	Car	Phe	Aen	LVS	Tle	Val	Asn	Phe	Phe	Asp	Gln	Lys	Pro	Glu	Glu
	Gru	Ser	FIIC	A3II	144					145	0 -		_		1455	5
	_		Asn	-1	744	, ,	Acn	Cor	Glu	T.em	His	Ser	Asp	Ile	Arq	Lys
20	Leu	Hle	Asn			Leu	Abii	Ser	146				1	1470)	•
				146	0	_			740	2 21	mh~	7 cn	Tle			His
	Asn	Lys	Met	Asp	Ile	Leu	Ser	Tyr	GIU -	GIU	1111	ASP	110	· vai	_,,	
			1 4 7	_				148	0				T40:	,		
	Lys	Ile	Leu	Lys	Glu	Ser	Val	Pro	Val	Gly	Thr	GIY	Asn	GIN	Leu	Val
25							7 / 01	_				150	U			
	Thr	Phe	0 Gln	Gly	Gln	Pro	Glu	Arg	Asp	Glu	Lys	Ile	Lys	Glu	Pro	Thr
	160	c				757	()				Tor	_				
	Lau	J T.611	Gly	Phe	His	Thr	Ala	Ser	Gly	Lys	Lys	Val	Lys	Ile	Ala	Lys
	neu	пеп	GIY	FIIC	152					153	0				153	5
		_	Leu	•	152	2127	Tare	λen	1.611	Dhe	Asp	Glu	Lvs	Glu	Gln	Gly
30	Glu	Ser	ьeu			Vai	пуз	Maii	154				2 -	155	ם	-
				154	0			_	134		- T	או א	Larg			Lvs
	Thr	Ser	Glu	Ile	Thr	Ser	Phe	ser	HIS	GIN	rrp	ALA	Dys		шси	<i>-,</i>
			155	-				156	0				120	,		
	Tyr	Arq	Glu	Ala	Cys	Lys	Asp	Leu	Glu	Leu	Ala	Cys	GIU	Thi	TTE	GIU
35							157	Ε				156	U			
-	Tle	Thr	· Ala	Ala	Pro	Lys	Cys	Lys	Glu	Met	Gln	Asn	Ser	Leu	Asn	Asn
	2 5 0	_				159	Λ				723	φ .				100
	Yer	Tare	. Asn	Leu	Val	Ser	Ile	Glu	Thr	· Val	. Val	Pro	Pro	Lys	Leu	Leu
					760	15				T P T	. U				101	_
4.0	0		Asn	Tou	Cve	Aro	Gln	Thr	Glu	Asr	Leu	Lys	Thr	Ser	Lys	Ser
40	Sei	Ast) Mbi			, ,,,,			162	5		-		163	0	
			_	162		T	17-1	Uio	Glu	Δer	. Val	Glu	Lvs	Glu	Thr	Ala
	Ilε	Phe			yaı	. гуз	val	164	, G10	L ASI			164	5		
			163	5	_		_	164				. D~c			Val	Tle
	Lys	s Sei	r Pro) Ala	a Thr	. Cys	Tyr	Thr	ASI	1 GII	1 561	. PIC	TYL	SCI	741	Ile
45		16	= A				165	5				100				
	Gli	ı Ası	a Ser	: Ala	a Leu	ı Ala	Phe	Туг	Thr	: Sei	r Cys	s Ser	Arg	гAа	IIIL	Ser
	166	: =				167	0				167	/5				100
	Val	Sei	r Glr	1 Thi	r Sex	Lev	Lev	Gli	. Ala	Ly:	s Lys	Trp	Leu	Arg	Glu	Gly
					168	35				16:	90				103	5
E 0	т 1.	n Dh	- J.C.	- G1:	, Glr	Pro	Glu	Arc	ı Ile	Ası	a Thi	c Ala	Asp	Tyr	· Val	Gly
50	110	5 PIN	c Asi				, 01		170	35			_	171	.0	
		_	_	170			. 20-		· Ner	. Se	r Th	- T14	A la	Glu	Asn	qaA
	As	n Ty			r GII	1 ASI	1 . ASI	1 261	. Asi				172	5		-
			17	15			_	172	٥٠	_ m_	- The	~ T.O.			Ser	Ser
	Ly	s As:	n Hi	s Le	u Se	r GT	т г.	- G11	n Asi	o in	L LY.	L	. 201	. Ast		Ser
55		17	30				173	35				174				
	Me	t Se	r As:	n Se	r Ty:	r Se	r Ty	: Hi:	s Se	r As	p Gl	u Va.	г туз	ASI	ASE	Ser
	27	1 E				17	50				17	55				1,0
	G1	ν Τυ	r Le	u Se	r Lv	s As	n Ly	s Le	u As	p Se	r Gl	y Ile	e Glu	ı Pro	val	L Leu
					17	6 5				17	70					, ,
60	T	a 7 a	n 17a	ום ן	11 Ag	n Gl	n Lv	a As	n Th	r Se	r Ph	e Se	r Ly	s Val	llle	e Ser
90				17	RΛ				17	85				1/2	, ,	
	-		. 7	, T	ינע דע מ	2 N ~	n 1/1	a ጥኒ፣	r Pr	ດີເກ	n Th	r Va	l As	n Gli	ı Ası	ıle
	As	n va	LI LIY	BAB	א אד	a no	** **	~ + 7							•	

		1795				1800)				1805			
	Cys Val	Glu (101	F				1020				
5	Ala Ala	Ile :		10	20				TDJJ					
	Pro Pro		-	~ 4 5				1850	,					,
10	Glu Thr						I N h "	•				70,0	,	
	Val Ile					188	()				700-	,		
	Ile Met				7 2	95				TAGA				
15	1890 His Asn 1905			7 (חדב				1313					
	Phe Ala		-	ヘコモ				1930	J					_
20	Ser Gly		4 -				1 444	~				エフン・	,	
20	Glu Thr		Asp I	le C		1 4 5	13				100	_		
	Val Ser	Ser	Ala A		7 (3	'/ L				1300	,			
25	Ser Val	Gln		1	ดอก				199	3				
	1985 Ser Glu	Ile	Glu A	Asp S	er Th	r Lys	Gln	Val 201	Phe	Ser	Lys	Val	Leu 201	Phe 5
2.0	Lys Ser		Glu F				202	Thr 5	Arg			203	•	
30	Ile Arg		Pro (204	. (1				204	_		
	Val Val 205	Asn	Ser S		20	155				200	,			
35	Gln Val	Ser		2	ለማለ				20/	_				200
	Glu Glu			200E				209	U				2.02	_
40	Thr Ser		Gln	Asn V			210	15				2	•	
	Asn Pro		_			791	วก				~			
	Glu Phe	e Lys	Leu		າ:	175				2 14	v			
45	Asn Ası				2150				213	3				
	2145 Asp Ly	s Gln	Gln	Leu \ 2165	/al L	eu Gl	y Thi	Lys 217	val	Ser	Lev	va]	. Glu 217	ı Asn 75
	Ile Hi	s Val	Leu	Gly 1	Lys G	lu Gl	n Ala 218	a Ser	Pro	Lys	Ası	1 Val	Lys 0	Met
50	Glu Il			Thr (Glu T	hr Ph	e Se:	r Ası	y Val	Pro	Val 220	L Lys	Thi	c Asn
	Ile Gl	٦.	. Сув		2	yr Se 215	r Ly:			222				
55	Thr Gl	u Ala	a Val	Glu	lle A 2230	la Ly	s Al	a Pho	e Met 22:	: Glu 35	. Asj	p Ası	Gl:	u Leu 224
	2225 Thr As			Leu	Pro S			22	r Hi: 50	s Ser			22	<i></i>
	Pro Gl	u Ası	n Glu 226	Glu	Met V	al Le	u Se	r As 65	n Se	r Arg	, Il	e Gl	у L y 70	s Arg
60	Arg Gl	ly Gl	u Pro	Leu	Ile I	eu Va 22	al Gl 280	y Gl	u Pr	o Sei	r Il 22	e Ly 85	s Ar	g Asn

		2201	Asn				2295					2300)			
		Ala	Ser	Lys	Ser	Thr	Pro	Asp	Gly	Thr	Ile 2315	Lys	qaA	Arg	Arg	Leu 232
5	2309		His	ni o	17-1	2310	I.en	Glu	Pro	Tle			Val	Pro	Phe	
					2325	:				2330)				2335)
	Thr	Thr	Lys	Glu 2340		Gln	Glu	Ile	Gln 2345	Asn	Pro	Asn	Phe	Thr 2350	Ala)	Pro
10	Gly	Gln	Glu	Phe	Leu	Ser	Lys	Ser	His	Leu	Tyr	Glu	His 2365	Leu	Thr	Leu
			2355 Ser			202	Ten	2360) V21	Ser	Glv	His			Tvr	Gln
		2370	1				2375	i				2380)			
	Val	Ser	Ala	Thr	Arg			Lys	Met	Arg	His	Leu -	Ile	Thr	Thr	G1y 240
15	238	5		_		2390) 77- 1	D	Dwo	Dho	2399		TAVE	Ser	His	
			Thr		2409	5				241)				241:	•
	His	Arg	Val	Glu	Gln	cys	Val	Arg	Asn	Ile	Asn	Leu	Glu	Glu	Asn	Arg
				2420)			•	2425			3	C ~ ~	2430		Tare
20			Gln 243					2440)				244	5		
	Ile	Asn	Asp	Asn	Glu	Ile	His	Gln	Phe	Asn	Lys	Asn	Asn	Ser	Asn	Gln
		245	0			_,	2455	,	~	a 3	a 3	2460		T.em	Aen	T.e.11
0.5			Ala	Val	Thr	Pne 2470		гав	CVS	GIU	247	5 GIU	FIO	пец	лър	248
25	246	5 Thr	Ser	T.011	Gln	Agn	Ala	Ara	asp	Ile			Met	Arg	Ile	Lys
					2485	5				249	0				249	>
	Lys	Lys	Gln	Arg 2500	Gln	Arg	Val	Phe	Pro 250	Gln	Pro	Gly	Ser	Leu 251	Tyr	Leu
30	Ala	Lys	Thr	Ser	Thr	Leu	Pro	Arg 2520	Ile	Ser	Leu	Lys	Ala 252	Ala 5	Val	Gly
	a 3	<i>C</i> 15	2519 Val	D*0	Cor	Δla	Cvs	Ser	His	Lvs	Gln	Leu			Tyr	Gly
		253	n				2535	5				254	0			
			Ъув	His	Cys			Ile	Asn	ser	Lys 255	ASN	Ala	GIU	261	256
35	254	5 Dha	His	Thr.	Glu	255	ህ ጥህን	Phe	Glv	Lvs			Leu	Trp	Thr	
					256	5				257	0				257	⊃
			Ile	258	0				258	5				259	U	
40	Gly	Lys	Ala	Gly	Lys	Glu	Glu	Phe	Tyr	Arg	Ala	Leu	Cys 260	Asp	Thr	Pro
	~1	** - 1	259 Asp	5	T	Len	Tla	260) Ara	Tle	Trn	Val			His	Tyr
	GIY	vai 261		PIO	гур	neu	261	5	n. 9			262	0			•
	Arc	Tru	Ile	Ile	Trp	Lys	Leu	Ala	Ala	Met	Glu	Сув	Ala	Phe	Pro	Lys
45	262	5				263	0				263	5				264
	Glu	Phe	Ala	Asn			Leu	Ser	Pro	Glu	Arg	Val	Leu	Leu	Gin	Leu
			_	_	264	5	01	T 1.0	N am	265		Δνα	Δτα	Ser	265 Ala	
			Arg	266	0				266	5				267	U	
50	Lys	Lys	: Ile	Met	Glu	Arg	Asp	Asp	Thr	Ala	Ala	Lys	Thr	Leu	Val	Leu
			267	5	_		_	268				-73	268		Thr	Cer
	-	269	L Ser				269	5				270	0			
	Sex	c Ası	ı Lys	Thr	Ser	Ser	Ala	Asp	Thr	Glr	Lys	Val	Ala	ı Ile	Ile	Glu
55	276	15				271	.0				271	.5				212
	Le	ı Thi	r Asp	Gly	Trp 272		Ala	. Val	ГЛЗ	Ala 273	ı Glr 30	ı Lev	As <u>r</u>) Pro	273	ь ьец 5
	Lei	וא נו	a Val	Leu	Lys	s Asr	Gly	Aro	Lev			Gly	, Glr	ı Lys		
				274	10				274	5				275	U	
60	Le	u Hi	s Gly		ı Glı	ı Lev	val			Pro	Ası) Ala	Cys	Thr	Pro	ьeu
	~ 1		27! a Pro	55	, ce	r T.01	1 Met	276 Te:		. T1	s Sei	c Ala	276 a Asi		Thr	. Arq
	G I	u A.⊾	a r:(וגיט ע	7 251	باندر			y "							_

	2770		2775		780	
	Pro Ala Arg 2785	Trp Tyr Thr		y Phe Phe P 2795	ro Asp Pro	Arg Pro 280
5	Phe Pro Leu	Pro Leu Ser 2805	Ser Leu Ph	e Ser Asp G 2810	ly Gly Asn	Val Gly 2815
	•	Val Ile Ile 2820	28	25	283	0
10	283		2840		2845	
	2850	Ala Lys Tyr	2855	2	860	
	2865	Lys Ile Gln 287	0	2875		288
15		Tyr Leu Pro 2885		2890		2895
		Asp Gly Ala 2900	29	05	291	0
20	291		2920		2925	
	2930	His Arg Gln	2935	2.	940	
	2945	Ile Arg Lys 2950	0	2955		296
25	-	Arg Asp Val 2965		2970		2975
		Lys Glu Lys 2980	29	85	2990	0
30	Ser Ser Asp	Leu Tyr Ser	Leu Leu Th	r Glu Gly L	ys Arg Tyr 3005	Arg Ile
	Tyr His Leu 3010	Ala Thr Ser	3015	3 (020	
	3025	Ala Ala Thr	٥	3035		304
35	_	Ile Leu Phe 3045		3050		3055
	_	Phe Leu Asp 3060	30	55	3070	כ
40	307		3080		3085	
	3090	Tyr Leu Ser	3095	3	100	
	=	Asp Leu Asn 3110 Asn Leu Gln	_			
45		3125		3130		3135
		Ala Gly Asp 3140	31	45	3150	D
50	315	-	3160		3165	_
	3170	Leu Cys Asn	3175	3	180	_
	3185	Asp Pro Lys	0	3195		320
55	-	Thr Ala Gln 3205		3210		3215
		Pro Asn Cys	32	25	323	0
60	323		3240		3245	
	Thr Ser Lys 3250	Ser Cys Lys	Gly Glu Ly 3255		sp Asp Gin	Lys Asn

	Cys 3265	_	Lys	Arg	Arg	Ala 3270		Asp	Phe	Leu	Ser 3275		Leu	Pro	Leu	Pro 328	
	Pro	Pro	Val	Ser		Ile		Thr	Phe			Pro	Ala	Ala	Gln	Lys	
5	_			_	3285	•	0	~	a1	3290		т	<i>a</i> 1	Thr	3295		
				3300)				3305	;				3310			
	Lys	Lys	Lys 3315		Leu	Asn	Ser	Pro 3320		Met	Thr	Pro	Phe 3325	Lys	Lys	Phe	
10	Asn		Ile	Ser	Leu	Leu		Ser		Ser	Ile			Glu	Glu	Leu	
	Ala	3330 Leu	Ile	Asn	Thr	Gln	3335 Ala		Leu	Ser	Gly	3340 Ser		Gly	Glu	Lys	
	3345	i				3350)				3355	;				336	
	Gln	Phe	Ile	Ser	Val 3365		Glu	Ser	Thr	Arg 3370		Ата	Pro	Thr	Ser 3375	ser	
15	Glu	Asp	Tyr		Arg		Lys	Arg		Cys		Thr	Ser	Leu	Ile		
	Glu	Gln	Glu	3380 Ser		Gln	Ala	Ser	3385 Thr		Glu	Cys	Glu	3390 Lys	Asn	Lys	
			3395		_,	-1.	~	3400		T 1.0			3405	5			
20	Gln	Asp 3410		Ile	Thr	Thr	Lys 3415		Tyr	TTE							
			(2)	INE	ORM	TION	FOR	SEÇ	ID	NO : 8	3:						
25		/;	1 01	OTTEN	ICE (HARA	משרט	ודפו	cs:								
25		(1				1048				3							
						ıclei											
						NESS (: li			:								
30			(1)	IOFC	200.												
		•				TYPE	: cE	ANG									
		()	.x) F	EATU)RE:												
						Y: C				ice							
35			(B)	LOC	CATIO	N: 2	29 MATI	.104 ON:	BRC	A2 (C	OMI3)						
		()	ci) S	SEQUE	ENCE	DESC	RIPI	rion:	SEÇ	מד נ	NO:8	s :					
40	GGTC	GCG	GA C	CTT	CTGA	AA CI	CAGGO	CGGC	GAC	GCG	GAGC	CGCT	GTGC	CA (CTGCT	rgcgcc	60
	TCTC	CTG	CGC C	TCG	GTG:	rc Ti	TTGC	CGGCG	GTO	GGT(CGCC TCCG	GCCG	GGAC	SAA (SCGTO AACTO	BAGGGG BCACCT	120 180
	CTG	GAGC	GA (TTA:	TTTA(CC AF	AGCAT	TGG	GGZ	ATA!	rcgt	AGGT	'AAA	TA	G CCT	TTA	237
																lle	
45														1			
	GGA	TCC	AAA	GAG	AGG	CCA	ACA	TTT	TTT	GAA	ATT	TTT	AAG	ACA	CGC	TGC	285
	Gly		Lys	Glu	Arg	Pro	Thr 10	Phe	Phe	Glu	Ile	Phe 15	ГЛя	Thr	Arg	Cys	
50		5															
	AAC	AAA	GCA	GAT	TTA	GGA	CCA	ATA	AGT	CTT	AAT	TGG	TTT	GAA	GAA	CTT	333
	Asn 20	Lys	Ala	Asp	Leu	GLy 25	Pro	He	Ser	Leu	Asn 30	Trp	Pne	GIU	Glu	35	
55	TCT	TCA	GAA	GCT	CCA	CCC	TAT	AAT	TCT	GAA	CCT	GCA	GAA	GAA	TCT	GAA Glu	381
	ser	ser	Glu	ALA	Pro 40	Pro	Tyr	ASN	ser	45	PIO	WIG	GIU	GIU	Ser 50	314	
											m		7 ~~	003	(13 B	N.C.C	420
60	CAT	AAA	AAC	AAC	AAT Agn	TAC Tvr	GAA Glu	Pro	AAC	CTA Leu	Phe	AAA Lvs	Thr	Pro	CAA Gln	Arg	429
30	1113	בי עם	NO II	55	11011	-1-	4		60	_04		-1-		65		_	

5	AAA Lys	CCA Pro	TCT Ser 70	TAT Tyr	AAT Asn	CAG Gln	CTG Leu	GCT Ala 75	TCA Ser	ACT Thr	CCA Pro	ATA Ile	ATA Ile 80	TTC Phe	AAA Lys	GAG Glu	477
5	CAA Gln	GGG Gly 85	CTG Leu	ACT Thr	CTG Leu	CCG Pro	CTG Leu 90	TAC Tyr	CAA Gln	TCT Ser	CCT Pro	GTA Val 95	AAA Lys	GAA Glu	TTA Leu	GAT Asp	525
10	AAA Lys 100	TTC Phe	AAA Lys	TTA Leu	GAC Asp	TTA Leu 105	GGA Gly	AGG Arg	AAT Asn	GTT Val	CCC Pro 110	AAT Asn	AGT Ser	AGA Arg	CAT His	AAA Lys 115	573
15				ACA Thr													621
20				CTA Leu 135													669
25				GTA Val													717
23				ACA Thr													765
30	ATT Ile 180	TCT Ser	GAA Glu	AGT Ser	CTA Leu	GGA Gly 185	GCT Ala	GAG Glu	GTG Val	GAT Asp	CCT Pro 190	GAT Asp	ATG Met	TCT Ser	TGG Trp	TCA Ser 195	813
35				GCT Ala													861
40	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
45	AAT Asn	GTG Val	AAA Lys 230	AGC Ser	TAT Tyr	TTT Phe	TCC Ser	AAT Asn 235	CAT His	GAT Asp	GAA Glu	AGT Ser	CTG Leu 240	AAG Lys	AAA Lys	AAT Asn	957
43				ATC Ile													1005
50				AGT Ser			Phe										1053
55				TGC Cys													1101
60				GTA Val 295													1149
	TTT	TCA	TTA	TGT	TTT	TCT	AAA	TGT	AGA	ACA	AAA	AAT	CTA	CAA	AAA	GTA	1197

	Phe	Ser	Leu 310	-	Phe	Ser	Lys	Cys 315	Arg	Thr	Lys	Asn	Leu 320	Gln	Lys	Val	
5																GAT Asp	1245
10												AAA Lys				GTA Val 355	1293
15		_										GAT Asp				GCA Ala	1341
20												ATC Ile					1389
20												ACC Thr					1437
25												CAT His 415					1485
30												ACA Thr					1533
35												CCA Pro					1581
40												ACA Thr					1629
10												GAC Asp					1077
45												GCT Ala 495					1725
50												TCA Ser					1773
55												CCA Pro					1821
60												CAT His					1869
00												GAT Asp					1917

CCA GCC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT Ile Ser Thr Leu Lys Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp Gln Lys Ser GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr TGT TCT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro Glu Ala Asp TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp Pro Lys Ser AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Cys His CCA GTA CAA CAC TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA

Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg
775 780 785

5	GAA Glu	TCT Ser 805	GAT Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	AAA Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
10	GAT Asp 820	GTA Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
10												AGA Arg					2781
15	AAC Asn	CAA Gln	AAC Asn	ACA Thr 855	AAT Asn	CTA Leu	AGA Arg	GTA Val	ATC Ile 860	CAA Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GAA Glu	ACT Thr	2829
20												TCT Ser					2877
25												GCT Ala 895					2925
30												GAA Glu					2973
30												ATG Met					3021
35	GAC Asp	ACA Thr	GGT Gly	GAT Asp 935	AAA Lys	CAA Gln	GCA Ala	ACC Thr	CAA Gln 940	GTG Val	TCA Ser	ATT Ile	AAA Lys	AAA Lys 945	GAT Asp	TTG Leu	3069
40												GTA Val					3117
45												ATC Ile 975					3165
50		Lys										GAC Asp					3213
30				Pro					Ser			GGT Gly		Phe			3261
55			naA					Leu				AAC Asn	Ile				3309
60		Met					Ile					CCT Pro					3357

5	TGT GTT Cys Val 1045	GAA Glu	ATT Ile	GTA Val	Asn	ACC Thr	TTG Leu	GCA Ala	TTA Leu	Asp	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
5	AGC AAG Ser Lys 1060	CCT Pro	CAG Gln	Ser	ATT Ile 1065	AAT Asn	ACT Thr	GTA Val	Ser	GCA Ala L070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
10	GTA GTT Val Val		Ser					Ser					Gln			3501
15	TTT TCC Phe Ser	Lys 1	Gln 1095	Asp	Phe	Asn	Ser	Asn 100	His	Asn	Leu	Thr	Pro 1105	Ser	Gln	3549
20		Glu 1110	Ile	Thr	Glu	Leu 1	Ser 1115	Thr	Ile	Leu	Glu	Glu 1120	Ser	Gly	Ser	3597
25	CAG TTT Gln Phe 1125	Glu	Phe	Thr	Gln 1	Phe 130	Arg	Lys	Pro	Ser	Tyr 1135	Ile	Leu	Gln	Lys	3645
	AGT ACA Ser Thr 1140	Phe	Glu	Val	Pro 1145	Glu	Asn	Gln	Met 1	Thr 1150	Ile	Leu	Lys	Thr 1	Thr 155	3693
30	TCT GAG Ser Glu	Glu	Cys	Arg 1160	Asp	Ala	Asp	Leu :	His 1165	Val	Ile	Met	Asn :	Ala 1170	Pro	3741
35	TCG ATT Ser Ile	Gly	Gln 1175	Val	Asp	Ser	Ser	Lys 180	Gln	Phe	Glu	Gly	Thr 1185	Val	Glu	3789
40		Arg 1190	Lys	Phe	Ala	Gly 1	Leu 195	Leu	Lys	Asn	Asp	Cys 1200	Asn	Lys	Ser	3837
45	GCT TCT Ala Ser 1205	Gly	Tyr	Leu	Thr	Asp 1210	Glu	Asn	Glu	Val	Gly L215	Phe	Arg	Gly	Phe	3885
	TAT TCT Tyr Ser 1220			Gly					Val					Leu		3933
50	AAA GCT Lys Ala	Val	Lys	Leu 1240	Phe	Ser	Asp	Ile	Glu L245	Asn	Ile	Ser	Glu :	Glu 1250	Thr	3981
55	TCT GCA Ser Ala	Glu	Val 1255	His	Pro	Ile	Ser	Leu 1260	Ser	Ser	Ser	Lys	Cys 1265	His	Asp	4029
60	TCT GTT Ser Val					Lys					Asn					4077
	AGT GAA	AAA	AAT	AAT	AAI.	TGC	CAA	CTG	ATA	TTA	CAA	AAT	AAT	ATT	GAA	4125

	Ser Glu 1285		Asn	Asn		Cys .290	Gln	Leu	Ile		Gln 1295	Asn	Asn	Ile	Glu	
5	ATG ACT Met Thr 1300	ACT Thr	GGC Gly	Thr	TTT Phe 305	GTT Val	GAA Glu	GAA Glu	Ile	ACT Thr 1310	GAA Glu	AAT Asn	TAC Tyr	Lys	AGA Arg 1315	4173
10	AAT ACT Asn Thr	GAA Glu	Asn	GAA Glu L320	GAT Asp	AAC Asn	AAA Lys	Tyr	ACT Thr 1325	GCT Ala	GCC Ala	AGT Ser	Arg	AAT Asn 1330	TCT Ser	4221
15	CAT AAC His Asn	Leu	GAA Glu 1335	TTT Phe	GAT Asp	GGC Gly	Ser	GAT Asp 1340	TCA Ser	AGT Ser	AAA Lys	Asn	GAT Asp 1345	ACT Thr	GTT Val	4269
20	TGF ATT Cys Ile	CAT His 1350	AAA Lys	Aap	GAA Glu	Thr	GAC Asp 1355	TTG Leu	CTA Leu	TTT Phe	Thr	GAT Asp 1360	CAG Gln	CAC His	AAC Asn	4317
20	ATA TGT Ile Cys 1365	Leu	AAA Lys	TTA Leu	Ser	GGC Gly L370	CAG Gln	TTT Phe	ATG Met	Lys	GAG Glu 1375	GGA Gly	AAC Asn	ACT Thr	CAG Gln	4365
25	ATT AAA Ile Lys 1380	GAA Glu	GAT Asp	Leu	TCA Ser 1385	GAT Asp	TTA Leu	ACT Thr	Phe	TTG Leu L390	GAA Glu	GTT Val	GCG Ala	Lys	GCT Ala .395	4413
30	CAA GAA Gln Glu	GCA Ala	Cys	CAT His 1400	GGT Gly	AAT Asn	ACT Thr	Ser	AAT Asn 1405	AAA Lys	GAA Glu	CAG Gln	Leu	ACT Thr 1410	GCT Ala	4461
35	ACT AAA Thr Lys	Thr	GAG Glu 1415	CAA Gln	AAT Asn	ATA Ile	Lys	GAT Asp 1420	TTT Phe	GAG Glu	ACT Thr	Ser	GAT Asp 1425	ACA Thr	TTT Phe	4509
4.0	TTT CAG Phe Gln	ACT Thr 1430	GCA Ala	AGT Ser	GGG Gly	Lys	AAT Asn L435	ATT Ile	AGT Ser	GTC Val	Ala	AAA Lys 1440	GAG Glu	TCA Ser	TTT Phe	4557
40	AAT AAA Asn Lys 1445	Ile	GTA Val	AAT Asn	Phe	TTT Phe 1450	GAT Asp	CAG Gln	AAA Lys	Pro	GAA Glu 1455	GAA Glu	TTG Leu	CAT His	AAC Asn	4605
45	TTT TCC Phe Ser 1460	TTA	AAT Asn	Ser	GAA Glu 1465	TTA Leu	CAT His	TCT Ser	Asp	ATA Ile 1470	AGA Arg	AAG Lys	AAC Asn	Lys	ATG Met L475	4653
50	GAC ATT Asp Ile		Ser					Asp					Lys			4701
55	AAA GAA Lys Glu	AGT Ser	GTC Val 1495	CCA Pro	GTT Val	GGT Gly	Thr	GGA Gly 1500	AAT Asn	CAA Gln	CTA Leu	Val	ACC Thr 1505	TTC Phe	CAG Gln	4749
~~	GGA CAA Gly Glr	CCC Pro	Glu	CGT Arg	GAT Asp	Glu	AAG Lys 1515	ATC Ile	AAA Lys	GAA Glu	Pro	ACT Thr 1520	CTG Leu	TTG Leu	GGT Gly	4797
60	TTT CAT	C ACA	GCT Ala	AGC Ser	GGG Gly	AAA Lys	AAA Lys	GTT Val	AAA Lys	ATT Ile	GCA Ala	AAG Lys	GAA Glu	TCT Ser	TTG Leu	4845

1525 1530 1535

5				AG CAA GGT ACT lu Gln Gly Thr 50	
10				CC CTA AAG TAC hr Leu Lys Tyr	
15		sp Leu Glu Leu		CC ATT GAG ATC hr Ile Glu Ile 1585	
		ys Lys Glu Met		TC AAT AAT GAT eu Asn Asn Asp 1600	
20				AG CTC TTA AGT ys Leu Leu Ser 1615	
25				CA AAA AGT ATC er Lys Ser Ile 30	
30	Lys Val Lys V	al His Glu Asn 1640	Val Glu Lys G		Ser Pro .650
35		yr Thr Asn Gln		CA GTC ATT GAA er Val Ile Glu 1665	
		he Tyr Thr Ser		AA ACT TCT GTG ys Thr Ser Val 1680	
40	Thr Ser Leu L 1685	eu Glu Ala Lys 1690	Lys Trp Leu Ai	GA GAA GGA ATA rg Glu Gly Ile 1695	Phe Asp
45	Gly Gln Pro G	lu Arg Ile Asn	Thr Ala Asp Ty	AT GTA GGA AAT yr Val Gly Asn 10	Tyr Leu
50				AA AAT GAC AAA lu Asn Asp Lys 1	
55	Leu Ser Glu L	•		AC AGT AGC ATG sn Ser Ser Met 1745	
		yr His Ser Asp		AT GAT TCA GGA sn Asp Ser Gly 1760	
60			Gly Ile Glu P	CA GTA TTG AAG ro Val Leu Lys 1775	

5	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
10	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
15	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
20	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
25	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
30	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
35	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
40	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
45	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955	6093
	GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1960 1965 1970	6141
50	GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985	6189
55	GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000	6237
60	GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015	6285

	WO SZIOSZE.	-
	GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035	6333
5	CCA GAA CAT TTA ATA TCC CAA AAA GGC TII ICA IAI AAI GIG GIA DAN Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050	6381
10	TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065	6429
15	ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe 2070 2075 2080	6477
20	GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095	6525
	CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115	6573
25	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
30	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
35	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160	6717
40	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
45	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
50	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2225	6909
55	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
60	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG	7053

		-
	Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	
5	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
10	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2300 2305	7149
15	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
20	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2355	7293
25	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
30	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
35	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
40	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
45	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
50	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT Asn Glu lle His Gln Pho Asn Lys Asn Asn Ser Asn Gln Ala Ala Ala 2455 2460 2465	7629
55	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2495 2490 2495	7725
60	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr	7773

	WO 33703104			-
	2500	2505	2510	2515
5	TCC ACT CTG Ser Thr Leu	CCT CGA ATC TCT CTG Pro Arg Ile Ser Leu 2520	AAA GCA GCA GTA GGA Lys Ala Ala Val Gly 2525	GGC CAA GTT 7821 Gly Gln Val 2530
10	Pro Ser Ala	2555	2540	2545
	CAT TGC ATA His Cys Ile 2550	AAA ATT AAC AGC AAA Lys Ile Asn Ser Lys 2555	Ash Ala Giu Sel File	. 0111 1110 11-1
15	Thr Glu Asp 2565	TAT TTT GGT AAG GAA Tyr Phe Gly Lys Glu 2570	ger Leu Trp III Gr	עבון בבי
20	Gln Leu Ala 2580	GAT GGT GGA TGG CTC Asp Gly Gly Trp Leu 2585	2590	2595
25	GGA AAA GAA Gly Lys Glu	A GAA TTT TAT AGG GCT 1 Glu Phe Tyr Arg Ala 2600	CTG TGT GAC ACT CC. Leu Cys Asp Thr Pr 2605	A GGT GTG GAT 8061 D Gly Val Asp 2610
30	CCA AAG CTT	r ATT TCT AGA ATT TGC 1 lle Ser Arg Ile Trp 2615	G GTT TAT AAT CAC TA O Val Tyr Asn His Ty 2620	T AGA TGG ATC 8109 r Arg Trp Ile 2625
	ATA TGG AAM Ile Trp Lyc 263	A CTG GCA GCT ATG GA s Leu Ala Ala Met Gl 0 263	I Cys Ala Phe Plo by	8 014 1110 1111
35	AAT AGA TG Asn Arg Cy 2645	C CTA AGC CCA GAA AG s Leu Ser Pro Glu Ar 2650	G GTG CTT CTT CAA CT g Val Leu Leu Gln Le 2655	A AAA TAC AGA 8205 u Lys Tyr Arg
40	TAT GAT AC Tyr Asp Th 2660	G GAA ATT GAT AGA AG r Glu Ile Asp Arg Se 2665	C AGA AGA TCG GCT AT r Arg Arg Ser Ala II 2670	A AAA AAG ATA 8253 e Lys Lys Ile 2675
45	ATG GAA AG Met Glu Ar	G GAT GAC ACA GCT GC G Asp Asp Thr Ala Al 2680	A AAA ACA CTT GTT C a Lys Thr Leu Val Le 2685	CC TGT GTT TCT 8301 Eu Cys Val Ser 2690
50	GAC ATA AT Asp Ile Il	TT TCA TTG AGC GCA AA Le Ser Leu Ser Ala As 2695	AT ATA TCT GAA ACT TO In Ile Ser Glu Thr So 2700	CT AGC AAT AAA 8349 er Ser Asn Lys 2705
	ACT AGT AG Thr Ser Se 273	GT GCA GAT ACC CAA AA Br Ala Asp Thr Gln Ly 10 27:	As All Ala lie lie G	Id Dea IIII III
55	Gly Trp Tr 2725	AT GCT GTT AAG GCC C yr Ala Val Lys Ala G 2730	2735	eu deu iide
60	TTA AAG A Leu Lys A 2740	AT GGC AGA CTG ACA G sn Gly Arg Leu Thr V 2745	TT GGT CAG AAG ATT A al Gly Gln Lys Ile I 2750	TT CTT CAT GGA 8493 le Leu His Gly 2755

5	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
10	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800	8637
15	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
20	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 2825 2830 2835	8733
25	TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala 2840 2845 2850	8781
	GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 2860 2865	8829
30	AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880	8877
35	TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895	8925
40	GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala 2900 2905 2910 2915	8973
45	TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn 2920 2925 2930	9021
50	CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu 2935 2940 2945	9069
50	ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser 2950 2955 2960	9117
55	AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys 2965 2970 2975	9165
60	AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Sei Asp 2980 2985 2990 2995	9213

	WO 22/07/04	
	TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010	9261
5	GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025	9309
10	GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040	9357
15	ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055	9405
20	TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile 3060 3065 3070 3075	9453
	GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090	9501
25	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
30	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115 3120	9597
35	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
40	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
4.5	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
45	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
50	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
55	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
60	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG	9981

	Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	•
5	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys 3255 3260 3265	10029
10	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	10077
15	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
	CCA CCA AGG AGT TCT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys 3300 3305 3310 3315	10173
20	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	10221
25	TCT CTT TTG GAA AGT AAT TCA ATA GCI GAC GAA GAA CTT GCA TTG ATA Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	10269
30	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3355 3360	10317
35	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	10365
	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3380 3385 3390 3395	10413
40	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	10461
45	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	10485
50	(2) INFORMATION FOR SEQ ID NO:9:	
55	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
60	(ii) MOLECULE TYPE: protein(v) FRAGMENT TYPE: internal(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:	

			Ile													
-			Cys		Lys .											
5			Leu 35					A()					73			
			Glu				E 5					טט				
10		Gln	Arg			70					10					
	Phe		Glu		25					90					,,,	
15			Asp	100					105					110		
	_		Lys 115					120					123			
		110	Ser				135					T#0				
20			Gln			150					155					100
			Ser		165					1/0					± , →	
25			His	100					185					100		
			Ser 195 Val					200					203			
							215					220				
30			Asn			230					233					
			Asn Arg		245					250					200	
35			Lys	260					265					210		
			275 Leu					7 X D					203			
40		200					295				Cys	Arg				Leu
	205	-	. Val			วาภ				Lys	. Lys					220
			a Asp		つつに		Lys	Ser	Lys	Asn	Gln				Lys	
45				240				Pro	Asn	1				Leu		Ser
	Ası			. Asn	Gln	Lys	Pro	360 Phe	Glu	Ser	Gly	Ser 380	Asp		Ile	Ser
50	_		u Val	. Val	Pro	Ser 390	Let	, ı Ala	Суз	Glu	1 Trp	Ser		Leu	Thr	Leu 400
		r Gl			405	/ Ala	. Glr			410	s Ile	Pro			410	
55				427	Glr	Asn			425	5				42(,	Glu
JJ			4 2 1	g Lys	s Lys			441)				77.	,		Arg
			r Se	r Le			45	5				40	•			Val
60		l As	n Ly			471)				42 / ⋅					480
	11	e Le	u Al	a Va	l Ly	s Gl	n Al	a Il	e Se	r Gl	y Th	r Se	r Pro	υ va.	r wre	a Ser

					485					490	_	_,	•	a 1	495	Desa
				500					505					Glu 510		
5	-		515					520					525	Asp		
	Phe	Lys 530	Lys	Glu	Thr	Glu	Ala 535	Ser	Glu	Ser	Gly	Leu 540	Glu	Ile	His	Thr
10	Val 545	Сув	Ser	Gln	Lys	Glu 550	Asp	Ser	Leu	CÀa	Pro 555	Asn	Leu	Ile	Asp	Asn 560
10	Gly	Ser	Trp	Pro	Ala 565	Thr	Thr	Thr	Gln	Asn 570	Ser	Val	Ala	Leu	Lys 575	Asn
	Ala	Gly	Leu	Ile 580	Ser	Thr	Leu	Lys	Lys 585	Lys	Thr	Asn	Lys	Phe 590	Ile	Tyr
15	Ala	Ile	His 595	Asp	Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605	Pro	ГÀЗ	Asp
		610					615					620		Ala		
20	625					630					635			Leu		640
					645					650				Glu	655	
				660					665					Cys 670		
25			675					680					685	Leu		
	-	690					695					700		Ile		
30	705		_			710					715			Glu		720
					725					730				Leu	735	
		_		740					745					Asp 750		
35			755					760					765	Ala		
		770					775					780		Leu		
40	785		-			790					795			Leu		800
					805					810				Pro	815	
	_			820					825					Asn 830		
45			835					840					845			Lys
		850					855					860		Lys		
50	865					870					875					Glu 880
					885					890				Val	895	
				900					905					His 910		
55			915					920					925			Val
		930)				935	i				940		Ser		
60	945					950					955					Lys 960 Ser
	Gln	His	3 11e	: гЛа	965		пел	гетХ	GID	970		. Lys	Jer	Asp	975	

	Leu	ASN	116	980	гув	TIE	Pro	GIU	985	ABN	Asn	Asp	ıyr	990	Asp	гу
5			995					100	0	Asn			100	5		
		1010	כ				101	5		Lys		1020)			
	1029	5		_		1030)	_	_	Ile	1039	5		_		104
10					104	5				Thr 1050)			_	1059	5
	-	_		106	ם כ				106					1076	0	
15			1075	5				1080	o "	Lys			1089	5		
		1090)				109	5		Asn		1100)			
20	1109	5		_		1110)			Leu Phe	1115	5				112
20					1125	5				1130 Glu	}	_			1135	5
	БСи	OIII	בענם	1140		FIIC	Oru	Val	1145		VOIT	GIII	Mec	1150		шес
25			1155	5				1160)	Ala			1165	5		
		1170)				1175	5		Ser		1180)			
2.0	1185	,				1190)			Gly	1195	,				120
30					1205	5				Asp 1210	•				1215	5
		_		1220)			_	1225					1230)	
35			1235	;			_	1240)	Ser Ile	_		1245	;		
	Giu	1250		SEI	MIA		1255		PIO	116		1260		Ser	ser	пуs
	Cys 1265		Asp	Ser	Val	Val 1270		Met	Phe	Lys	Ile 1275		Asn	His	Asn	Asp 128
40	_				1285	;			-	Cys 1290					1295	;
				1300)				1305					1310)	
45			1315	;				1320)	Asn	_	_	1325			
		1330	1				1335	,		Gly		1340				
50	1345					7750		_			1355					136
30					1365	i				Gly 1370 Asp				_	1375	, -
				1380)				1385	_				1390)	
55			1395	i				1400)	Ile			1405			
		1410)				1415	i				1420				
C 0	1425	i				1430)			Lys	1435					144
60	GIU	ser	rne	Asn	Lys 1445		val	Asn	Pne	Phe 1450	_	GIN	гàв	Pro	Glu 1455	
	Leu	His	Asn	Phe	Ser	Leu	Asn	Ser	Glu	Leu	His	Ser	Asp	Ile	Arg	Lys

				146					146					147		
	Asn	Lys	Met 147		Ile	Leu	Ser	Tyr 148		Glu	Thr	Asp	11e		Lys	Hi
5	Lys	Ile 149	Leu		Glu	Ser	Val	Pro		Gly	Thr	Gly 150	Asn		Leu	. Va
	Thr 150	Phe	Gln	Gly	Gln	Pro 151	Glu		Asp	Glu	Lys 151	Ile		Glu	Pro	Th:
10	Leu	Leu	Gly	Phe	His 152		Ala	Ser	Gly	Lys 153		Val	Lys	Ile	Ala 153	Ly
			Leu	154	0				154	5				155	0	
			Glu 155	5				156	0				156	5		_
15		157	-				157	5				158	0			
	158	5	Ala			159	0				159	5				160
20			Asn		160	5				1610	0			-	161	5
			Asn -	1620)				162	5				163	0	
25			Leu 1635	5				164	0				164	5		
25		165					165	5				166	0			
	1665	5	Ser			1670)				1679	5				168
30			Gln Asp		1685	5				1690)				1699	5
			Leu	1700)				1705	5			_	1710)	_
35			1715 His	5				1720)				1725	5		_
•		1730					1735	5				1740)			
	1745	5	Leu			1750)				1755	5				176
40			Val		1765	;				1770)				1775	5
			Lys	1780)				1785	;				1790)	
45			1795 Glu	5				1800)				1805	5		
		1810					1815	;				1820)			
	1825	5	Ala			1830)				1835	;				184
50			Ile		1845					1850)				1855	5
			Lys	1860)				1865	;				1870)	
55			1875 Ala	5				1880)				1885	•		_
		1890					1895	;				1900)			
	1905	5				1910)				1915					192
60			Asp		1925	,				1930)				1935	;
	Ser	Gly	Leu	Glu 1940		Val	Ser	Lys	Ile 1945		Pro	Сув		Val 1950		Leu

	Glu	Thr	Ser 195	qe <i>A</i>	Ile	Сув	Lys	Cys 196		Ile	Gly	Lys	Leu 196		Lys	Ser
5		197	0	Ala			197	5				198	0			
	198	5		Val		1990	כ				199	5				200
				Glu	2009	5				2010)				201	5
10	-			Glu 2020)		_		202	5				2030)	
		_	203					2040)				204	5		
15		205	0	Ser			205	5				206	0			
	206	5		Ile		2070)				207	5		_		208
20				Asp	2085	5				2090)				2095	5
20			_	Gln 2100)			_	2109	5		-		2110)	
			2115					2120)				2125	5		
25		213	3	Leu			213	5				214	ַ			
	2145	5		Ser		2150)				215	5				216
2.0	-	_		Gln	2165	5		_		2170)				2175	5
30				Leu 2180)	_			2185	5				2190)	
			2199					2200)				2205	5		
35		2210)	Сув			2219	5				2220)			
	2225	5		Val Lys		2230)				2235	5				224
4.0		_		_	2245	5				2250)				2255	;
40				Glu 2260 Pro)				2265	5				2270)	
	_	-	2275					2280)				2285	5		
45		229)	Lys		_	229	5				2300)			
	2309	5		His		2310)	_	_		2315	5	_			232
50				Glu	2325	5				2330)	_			2335	;
30			_	2340 Phe)				2345	5				2350)	
	_		235			·		2360)				2369	5		
55		237	0	Thr			2375	5				2380)			
	238	5		Lys	_	2390)	-		_	2399	5				240
60	_			Glu	2409	5				2410)		_		2415	i
00				2420 Asn)				2425	5				2430)	
	GIII	пÃв	GIII	Well	116	voh	GTA	1172	GIY	Ser	vah	rsp	261	- J 0	V211	-y -

			2439					244					2445			
	Ile	Asn 2450	Asp	Asn	Glu	Ile	His 245		Phe	Asn	Lys	Asn 2460		Ser	Asn	Glr
5	Ala 2469		Ala	Val	Thr	Phe 2470		ГÀв	Сув	Glu	Glu 2479		Pro	Leu	Asp	Let 248
	Ile	Thr	Ser	Leu	Gln 2489		Ala	Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 249	_
10	Lys	Lys	Gln	Arg 2500		Arg	Val	Phe	Pro 250		Pro	Gly	Ser	Leu 2510	_	Let
	Ala	Lys	Thr 2519		Thr	Leu	Pro	Arg 2520		Ser	Leu	Lys	Ala 2525		Val	Gly
	_	2530	-				2535	5				2540)			_
15	2545	5	Lys			2550)				2555	5				256
			His		2565	5				2570)				2579	5
20	-	_	Ile	2580)		_	_	2589	5				2590)	
	_	_	Ala 2595	5	-			2600) `	_			2605	5		
	Gly	Val 2610	Asp	Pro	Lys	Leu	11e 2619		Arg	Ile	Trp	Va1		Asn	His	Tyr
25	Arg 2629	Trp	Ile	Ile	Trp	Lys 2630	Leu		Ala	Met	Glu 2635	Cys		Phe	Pro	Lys 264
			Ala		2645	5				2650)				2655	5
30			Arg	2660)				2665	5				2670)	
			11e 2675	;				2680)				2685	i		
		2690					2695	i				2700)			
35	2705	5	Lys			2710)	_			2715					272
			Asp	_	2725	, ·			_	2730)				2735	5
40			Val Gly	2740)				2745	5				2750)	
			2755 Pro	<u>;</u>				2760)				2765	,		
4.5		2770)				2775	j				2780)			
45	2785	;	Arg			2790)				2795					280
			Leu -		2805	5				2810)				2815	5
50			Asp	2820)				2825	5				2830)	
	_		Ser 2835	5				2840)				2845	;		
	-	2850			_		2855	;				2860)			
55	2865	5	Thr	_		2870)				2875	i				288
		_	Pro	-	2889	5				2890)				2895	5
60			Gln	2900)				2905	5				2910)	
	Asp	PTO	Ala 291		тел	GIU		Tyr		ser	GIU		G1n 2925		мгд	wra

	Leu Asn 293		His	Arg	Gln	Met 293		Asn	Asp	Lys	Lys 294		Ala	Gln	Ile
5	Gln Leu	Glu	Ile	Arg	Lys 2950		Met	Glu	Ser	Ala 295		Gln	ГÀа	Glu	Gln 296
5	2945 Gly Leu	Ser	Arg	Asp			Thr	Val	Trp			Arg	Ile	Val	
	_		_	296	5				297	ס ֿ		_		297	5
	Tyr Ser	_	2980)	_	_		298	5				299	0	
10	Ser Ser	2995	5				3000)				300	5		
	Tyr His	0				301	5				302	0			
15	Ile Gln	Leu	Ala	Ala	Thr 3030		Lys	Thr	GIn	Tyr 303		Gln	Leu	Pro	Val 304
	Ser Asp	Glu	Ile	Leu 3049	Phe		Ile	Tyr	Gln 3050	Pro		Glu	Pro	Leu 305	His
	Phe Ser	-	Phe 3060	Leu		Pro	Asp	Phe 3065	Gln		Ser	Сув	Ser 307	Glu	
20	Asp Leu	Ile 3075	_	Phe	Val	Val	Ser 3080		Val	Lys	Lys	Thr 3089	_	Leu	Ala
	Pro Phe		Tyr	Leu	Ser	Asp 3099		Cys	Tyr	Asn	Leu 310		Ala	Ile	Lys
25	Phe Trp	-	Asp	Leu	Asn 3110	Glu		Ile	Ile	Lys 311	Pro		Met	Leu	Ile 312
	Ala Ala	Ser	Asn			Trp	Arg	Pro			Lys	Ser	Gly		
	Thr Leu	Phe	Ala	3125 Gly		Phe	Ser	Val	3130 Phe		Ala	Ser	Pro	3135 Lys	
30	Olean Miles		3140		mb	Dh.		3145		T	.	mb	3150		3
30	Gly His	3155					3160)				3165	5		
	Ile Asp		Leu	Cys	Asn	Glu 3175		Glu	Asn	Lys	Leu 3180		His	Ile	Leu
35	His Ala 3185	Asn	Asp	Pro	Lys 3190	-	Ser	Thr	Pro	Thr 3199	_	Asp	Сув	Thr	Ser 320
	Gly Pro	Tyr	Thr	Ala 3205		Ile	Ile	Pro	Gly 3210		Gly	Asn	Lys	Leu 3215	Leu
	Met Ser		Pro 3220	Asn		Glu	Ile	Tyr 3225	Tyr		Ser	Pro	Leu 3230	Ser	
40	Cys Met				Lys	Ser	Val			Pro	Val	Ser			Met
	Thr Ser	3235		Cua	T	C111	3240		~1.,	T10	A a n	3245		Tua	λαπ
	325	0		-	_	3255	5	_			3260)			
45	Cys Lys 3265	Lys	Arg	Arg	Ala 3270	Leu	qaA	Phe	Leu	Ser 3275	Arg	Leu	Pro	Leu	Pro 328
	Pro Pro				Ile					Ser					Lys
	Ala Phe		Pro 3300		Arg	Ser	Суз	Gly 3305		Lys	Tyr	Glu	Thr 3310		Ile
50	Lys Lys		Glu		Asn	<i>3</i> er	Pro 3320	Gln		Thr	Pro	Phe 3325	ГЛЗ		Phe
	Asn Glu 333		Ser	Leu	Leu .	Glu 3335		Asn	Ser	Ile	Ala 3340	Asp		Glu	Leu
	Ala Leu	Ile	Asn	Thr			Leu	Leu	Ser	_		Thr	Gly	Glu	_
55	3345 Gln Phe	Ile	Ser	Val	3350 Ser		Ser	Thr	Arg	3355 Thr		Pro	Thr	Ser	336 Ser
	Glu Nas	TT-1	Len	3365		Tazo	n-~	7.~~	3370		ጥሎ∽	Ce-	Leu	3375	
	Glu Asp		3380)				3385	,				3390)	
60	Glu Gln	Glu 3395		Ser	Gln	Ala	Ser 3400		Glu	Glu	Сув	Glu 3405		Asn	Lys
	Gln Asp			Thr	Thr	Lys			Ile						

3410 3415

5	(2) INFORMATION FOR SEQ ID NO:10:	
J	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 10485 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double	
10	(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: CDNA (ix) FEATURE:	
15	(A) NAME/KEY: Coding Sequence (B) LOCATION: 22910482 (D) OTHER INFORMATION: BRCA2 (OMI4)	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: ::	
	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC TCTGCTGCGC CTCGGGTGTC TTTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG ACAGATTTGT GACCGGCGCG GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAA ATG CCT ATT	60 120 180 237
25	Met Pro Ile	23,
30	GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys 5 10 15	285
35	AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu 20 25 30 35	333
33	TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu 40 45 50	381
40	CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg 55 60 65	429
45	AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu 70 75 80	477
50	CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp 85 90 95	525
c c	AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys 100 105 110 115	573
55	AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT TCC Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser 120 125 130	621
60	TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln 135 140 145	669

5	TGT Cys	ACA Thr	CAT His 150	GTA Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	GAT Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	TGT Cys	GGG Gly	AGT Ser	717
10	TTG Leu	TTT Phe 165	CAT His	ACA Thr	CCA Pro	AAG Lys	TTT Phe 170	GTG Val	AAG Lys	GGT Gly	CGT Arg	CAG Gln 175	ACA Thr	CCA Pro	AAA Lys	CAT His	765
10												GAT Asp					813
15	AGT Ser	TCT Ser	TTA Leu	GCT Ala	ACA Thr 200	CCA Pro	CCC Pro	ACC Thr	CTT Leu	AGT Ser 205	TCT Ser	ACT Thr	GTG Val	CTC Leu	ATA Ile 210	GTC Val	861
20												CAT His					909
25												AGT Ser					957
20												AAC Asn 255					1005
30	GAA Glu 260	GCT Ala	GCA Ala	AGT Ser	CAT His	GGA Gly 265	TTT Phe	GGA Gly	AAA Lys	ACA Thr	TCA Ser 270	GGG Gly	AAT Asn	TCA Ser	TTT Phe	AAA Lys 275	1053
35	GTA Val	AAT Asn	AGC Ser	TGC Cys	AAA Lys 280	GAC Asp	CAC His	ATT Ile	GGA Gly	AAG Lys 285	TCA Ser	ATG Met	CCA Pro	AAT Asn	GTC Val 290	CTA Leu	1101
40												TCT Ser					1149
45												AAT Asn					1197
50	AGA Arg	ACT Thr 325	AGC Ser	AAG Lys	ACT Thr	AGG Arg	AAA Lys	AAA Lys	ATT Ile	TTC Phe	CAT His	GAA Glu 335	GCA Ala	AAC Asn	GCT Ala	GAT Asp	1245
30	GAA Glu 340	TGT Cys	GAA Glu	AAA Lys	TCT Ser	AAA Lys 345	AAC Asn	CAA Gln	GTG Val	AAA Lys	GAA Glu 350	AAA Lys	TAC Tyr	TCA Ser	TTT Phe	GTA Val 355	1293
55	TCT Ser	GAA Glu	GTG Val	GAA Glu	CCA Pro 360	AAT Asn	GAT Asp	ACT Thr	GAT Asp	CCA Pro 365	TTA Leu	GAT Asp	TCA Ser	AAT Asn	GTA Val 370	GCA Ala	1341
60												ATC Ile					1389

_	GTA Val	CCG Pro	TCT Ser 390	TTG Leu	GCC Ala	TGT Cys	GAA Glu	TGG Trp 395	TCT Ser	CAA Gln	CTA Leu	ACC Thr	CTT Leu 400	TCA Ser	GGT Gly	CTA Leu	1437
5	AAT Asn	GGA Gly 405	GCC Ala	CAG Gln	ATG Met	GAG Glu	AAA Lys 410	ATA Ile	CCC Pro	CTA Leu	TTG Leu	CAT His 415	ATT Ile	TCT Ser	TCA Ser	TGT Cys	1485
10	GAC Asp 420	CAA Gln	AAT Asn	ATT Ile	TCA Ser	GAA Glu 425	AAA Lys	GAC Asp	CTA Leu	TTA Leu	GAC Asp 430	ACA Thr	GAG Glu	AAC Asn	AAA Lys	AGA Arg 435	1533
15	AAG Lys	AAA Lys	GAT Asp	TTT Phe	CTT Leu 440	ACT Thr	TCA Ser	GAG Glu	AAT Asn	TCT Ser 445	TTG Leu	CCA Pro	CGT Arg	ATT Ile	TCT Ser 450	AGC Ser	1581
20	Leu	Pro	Lys	Ser 455	Glu	Lys	Pro	Leu	Asn 460	Glu	Glu	Thr	GTG Val	Val 465	Asn	Lys	1629
25	AGA Arg	GAT Asp	GAA Glu 470	GAG Glu	CAG Gln	CAT His	CTT Leu	GAA Glu 475	TCT Ser	CAT His	ACA Thr	GAC Asp	TGC Cys 480	ATT Ile	CTT Leu	GCA Ala	1677
	Val	Lys 485	Gln	Ala	Ile	Ser	Gly 490	Thr	Ser	Pro	Val	Ala 495	TCT Ser	Ser	Phe	Gln	1725
30	Gly 500	Ile	Lys	Lys	Ser	Ile 505	Phe	Arg	Ile	Arg	Glu 510	Ser	CCT Pro	Lys	Glu	Thr 515	1773
35	Phe	Asn	Ala	Ser	Phe 520	Ser	Gly	His	Met	Thr 525	Asp	Pro	AAC Asn	Phe	Lys 530	Lys	1821
40	GAA Glu	ACT Thr	GAA Glu	GCC Ala 535	TCT Ser	GAA Glu	AGT Ser	GGA Gly	CTG Leu 540	GAA Glu	ATA Ile	CAT His	ACT Thr	GTT Val 545	TGC Cys	TCA Ser	1869
45	Gln	ГÀа	Glu 550	Asp	Ser	Leu	Cys	Pro 555	Asn	Leu	Ile	Asp	AAT Asn 560	Gly	Ser	Trp	1917
	Pro	Ala 565	Thr	Thr	Thr	Gln	Asn 570	Ser	Val	Ala	Leu	Lys 575	AAT Asn	Ala	Gly	Leu	1965
50	11e 580	Ser	Thr	Leu	Lys	Lys 585	Lys	Thr	Asn	Lys	Phe 590	Ile	TAT Tyr	Ala	Ile	His 595	2013
55	Asp	Glu	Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605	Pro	Lys	GAC Asp	Gln	Буз 610	Ser	2061
60	GAA Glu	CTA Leu	ATT lle	AAC Asn 615	Сув	TCA Ser	GCC	CAG Gln	Phe 620	Glu	GCA Ala	AAT Asn	GCT Ala	TTT Phe 625	GAA Glu	GCA Ala	2109
	CCA	CTI	ACA	TTT	GCA	raa .	GCI	GAT	TCA	GGI	TTA	TTG	CAT	TCT	TCT	GTG	2157

	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640	Ser	Ser	Val	
5	AAA Lys	AGA Arg 645	AGC Ser	TGT Cys	TCA Ser	CAG Gln	AAT Asn 650	GAT Asp	TCT Ser	GAA Glu	GAA Glu	CCA Pro 655	ACT Thr	TTG Leu	TCC Ser	TTA Leu	2205
10	ACT Thr 660	AGC Ser	TCT Ser	TTT Phe	GGG Gly	ACA Thr 665	ATT Ile	CTG Leu	AGG Arg	AAA Lys	TGT Cys 670	TCT Ser	AGA Arg	AAT Asn	GAA Glu	ACA Thr 675	2253
15	TGT Cys	TCT Ser	AAT Asn	AAT Asn	ACA Thr 680	GTA Val	ATC Ile	TCT Ser	CAG Gln	GAT Asp 685	CTT Leu	GAT Asp	TAT Tyr	AAA Lys	GAA Glu 690	GCA Ala	2301
0.0	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
20	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
25	rya AAA	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAĞ Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
30	CCA Pro 740	GTA Val	CAA Gln	CAT His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
35	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
4.0	ACT Thr	CCT Pro	ACT Thr	TCC Ser 775	AAG Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	AAC Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	TCT Ser	AGA Arg	2589
40	GGC Gly	AAA Lys	GAA Glu 790	TCA Ser	TAC Tyr	AA. Lys	ATG Met	TCA Ser 795	GAC Asp	AAG Lys	CTC Leu	AAA Lys	GGT Gly 800	AAC Asn	AAT Asn	TAT Tyr	2637
45	GAA Glu	TCT Ser 805	Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	Lys	AAT Asn	ATT Ile	CCC	ATG Met 815	GAA Glu	L ys	AAT Asn	CAA Gln	2685
50	GAT Asp 820	Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
55	CCT Pro	GAA Glu	AAA Lys	TAC	ATG Met 840	Arg	GTA Val	GCA Ala	TCA Ser	Pro 845	Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	TTC Phe	2781
	AAC Asn	CAA	A AAC	ACA Thr	Asn	CTA Leu	AGA Arg	GTA Val	ATC Ile	Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GAA Glu	ACT	2829
60	ACT Thr	TCI Sei	A ATT	TCF Ser	AAA Lys	ATA	ACT Thr	GTC Val	AAT Asi	CCA	GAC Asp	TCT Ser	GAA Glu	GAA Glu	CTT Leu	TTC	2877

PCT/US98/16905 WO 99/09164

	WO 99/091	64												P	L 1/US	60/10702
		870					875					880				
5	TCA GAC Ser Asp 885	AAT Asn	GAG Glu	AAT Asn	AAT Asn	TTT Phe 890	GTC Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	AAT Asn	GAA Glu	AGG Arg	AAT Asn	2925
10	AAT CTT Asn Leu 900	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
1.5	TGT GTA Cys Val	AAC Asn	GAA Glu	CCC Pro 920	ATT Ile	TTC Phe	AAG Lys	AAC Asn	TCT Ser 925	ACC Thr	ATG Met	GTT Val	TTA Leu	TAT Tyr 930	GGA Gly	3021
15	GAC ACA Asp Thr	GGT Gly	GAT Asp 935	AAA Lys	CAA Gln	GCA Ala	ACC Thr	CAA Gln 940	GTG Val	TCA Ser	ATT Ile	AAA Lys	AAA Lys 945	GAT Asp	TTG Leu	3069
20	GTT TAT Val Tyr	GTT Val 950	CTT Leu	GCA Ala	GAG Glu	GAG Glu	AAC Asn 955	AAA Lys	AAT Asn	AGT Ser	GTA Val	AAG Lys 960	CAG Gln	CAT His	ATA Ile	3117
25	AAA ATG Lys Met 965	ACT Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	TTA Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	AAT Asn	ATA Ile	3165
30	GAT AAA Asp Lys 980	ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	AAT Asn	AAT Asn	GAT Asp	TAC Tyr	ATG Met 990	AAC Asn	rys Tys	TGG Trp	GCA Ala	GGA Gly 995	3213
2.5	CTC TTA Leu Leu	GGT Gly	Pro	ATT Ile 1000	Ser	TAA Asn	CAC His	Ser	TTT Phe 1005	GGA Gly	GGT Gly	AGC Ser	Phe	AGA Arg 1010	ACA Thr	3261
35	GCT TCA Ala Ser	Asn	AAG Lys 1015	GAA Glu	ATC Ile	AAG Lys	Leu	TCT Ser 1020	GAA Glu	CAT His	AAC Asn	Ile	AAG Lys 1025	AAG Lys	AGC Ser	3309
40	AAA ATG Lys Met	TTC Phe 1030	Phe	AAA Lys	GAT Asp	Ile	GAA Glu 1035	GAA Glu	CAA Gln	TAT Tyr	Pro	ACT Thr 1040	AGT Ser	TTA Leu	GCT Ala	3357
45	TGT GTT Cys Val 1045	Glu	ATT Ile	GTA Val	Asn	ACC Thr 1050	TTG Leu	GCA Ala	TTA Leu	Asp	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
50	AGC AAG Ser Lys 1060	CCT Pro	CAG Gln	Ser	ATT Ile 1065	. sn	ACT Thr	GTA Val	Ser	GCA Ala 1070	His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
	GTA GTT Val Val	GTI Val	TCT Ser	GAT Asp 1080	Сув	AAA Lys	AAT Asn	AGT Ser	CAT His 1085	Ile	ACC Thr	CCT Pro	Gln	ATG Met 1090	Leu	3501
55	TTT TCC Phe Ser	AAC Lys	CAG Gln 1095	Asp	TTT Phe	AAT Asn	TCA Ser	AAC Asn 1100	His	AAT Asn	TTA Leu	Thr	CCT Pro 1105	Ser	CAA Gln	3549
60	AAG GC	A GAJ a Glu 1110	ı Ile	ACA Thr	GAA	CTI Leu	TCI Ser 1115	Thr	TATA	TTA Leu	GAA Glu	GAA Glu 1120	Ser	GGA Gly	AGT Ser	3597

5	CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135	3645
	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150	3693
10	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
15	TCG ATT GGT CAG GTA GAC AGC AGC CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
20	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA ATT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
25	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
30	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
35	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
40	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
45	AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310	4173
50	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
55	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
60	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317

	-	-
_	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 1370 1375	4365
5	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
10	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
15	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
20	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr A1a Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605
25	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475	4653
30	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
35	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
40	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
	TTT CAT ACA GCT AGC GGG AHA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
45	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1550 1555	4893
50	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570	4941
55	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
60	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT	5085

		-
	Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	
5	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1635	5133
10	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
15	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
20	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
25	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710	5373
30	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
35	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
40	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
45	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
50	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	566±
55	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
60	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile	5805

PCT/US98/16905 WO 99/09164 ARA ARA GTG ARA GAC ATA TTT ACA GAC AGT TTC AGT ARA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn TCA TCI GCT TTC TCT GGA. TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser

ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT

Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe
2070 2075 2080

GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA
Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg
2085 2090 2095

5	CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115	6573
	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
10	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
15	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2160	6717
20	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
25	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
30	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
35	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
40	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
45	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	7053
	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile L; Arg Asn Leu Leu Asn 2280 2285 2290	7101
50	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305	7149
55	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
60	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245

	W6727672	
	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2355	7293
5	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
10	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
15	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
20	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AC AGA CAA AAG CAA Glu Glu Glu Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
25	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
30	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT ABN Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
35	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
40	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2510 2515	7773
45	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
50	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 . 2540 2545	7869
55	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2560	7917
60	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT	8013

	Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595	
5	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610	8061
10	CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625	8109
15	ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2640	8157
	AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2005	8205
20	TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2675	8253
25	ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 2685 2690	8301
30	GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705	8349
35	ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720	8397
	GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val 2725 2730 2735	8445
40	TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755	8493
45	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
50	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
55	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800	8637
- -	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
60	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser	8733

	2020	2825	2830	2835
5	TCT GGA TTA Ser Gly Leu	mag ama more CCC	AAT GAA AGA GAG GAA GAA Asn Glu Arg Glu Glu Glu 2845	AAG GAA GCA 8781 Lys Glu Ala 2850
10	Ala Lys Tyr	GTG GAG GCC CAA Val Glu Ala Gln 2855	CAA AAG AGA CTA GAA GCC Gln Lys Arg Leu Glu Ala 2860	TTA TTC ACT 8829 Leu Phe Thr 2865
	AAA ATT CAG Lys Ile Gln 2870	Glu Glu Phe Glu	GAA CAT GAA GAA AAC ACA Glu His Glu Glu Asn Thi 875 2880	INI Dys FIO
15	TAT TTA CCA Tyr Leu Pro 2885	TCA CGT GCA CTA Ser Arg Ala Leu 2890	ACA AGA CAG CAA GTT CG Thr Arg Gln Gln Val Arg 2895	GCT TTG CAA 8925 3 Ala Leu Gln
20	GAT GGT GCA Asp Gly Ala 2900	GAG CTT TAT GAA Glu Leu Tyr Glu 2905	GCA GTG AAG AAT GCA GC Ala Val Lys Asn Ala Ala 2910	A GAC CCA GCT 8973 a Asp Pro Ala 2915
25	TAC CTT GAG Tyr Leu Glu	GGT TAT TTC AGT Gly Tyr Phe Ser 2920	GAA GAG CAG TTA AGA GCG Glu Glu Gln Leu Arg Al 2925	C TTG AAT AAT 9021 a Leu Asn Asn 2930
30	CAC AGG CAA His Arg Glr	A ATG TTG AAT GAT Met Leu Asn Asp 2935	AAG AAA CAA GCT CAG AT Lys Lys Gln Ala Gln Il 2940	C CAG TTG GAA 9069 e Gln Leu Glu 2945
	ATT AGG AAG Ile Arg Lys 2950	Ala Met Glu Ser	GCT GAA CAA AAG GAA CA Ala Glu Gln Lys Glu Gl 2955 296	n Gry neu ber
35	AGG GAT GTO Arg Asp Val 2965	C ACA ACC GTG TGG 1 Thr Thr Val Trp 2970	AAG TTG CGT ATT GTA AG Lys Leu Arg Ile Val Se 2975	C TAT TCA AAA 9165 r Tyr Ser Lys
40	AAA GAA AA Lys Glu Ly 2980	A GAT TCA GTT ATA s Asp Ser Val Ile 2985	CTG AGT ATT TGG CGT CC Leu Ser Ile Trp Arg Pr 2990	A TCA TCA GAT 9213 o Ser Ser Asp 2995
45	TTA TAT TC Leu Tyr Se	T CTG TTA ACA GAA r Leu Leu Thr Glu 3000	GGA AAG AGA TAC AGA AT Gly Lys Arg Tyr Arg Il 3005	T TAT CAT CTT 9261 e Tyr His Leu 3010
50	GCA ACT TO Ala Thr Se	A AAA TCT AAA AGT r Lys Ser Lys Ser 3015	AAA TCT JAA AGA GCT AF Lys Ser Glu Arg Ala As 3020	AC ATA CAG TTA 9309 sn lle Gln Leu 3025
	GCA GCG AC Ala Ala Th 303	ir Lys Lys Thr Gln	TAT CAA CAA CTA CCG G Tyr Gln Gln Leu Pro V 3035 304	al Ser Asp Glu
55	ATT TTA TI Ile Leu Pr 3045	TT CAG ATT TAC CAG ne Gln Ile Tyr Glr 3050	G CCA CGG GAG CCC CTT C n Pro Arg Glu Pro Leu H 3055	AC TTC AGC AAA 9405 is Phe Ser Lys
60	TTT TTA GA Phe Leu As 3060	AT CCA GAC TTT CAC sp Pro Asp Phe Gli 3065	G CCA TCT TGT TCT GAG G n Pro Ser Cys Ser Glu V 3070	TG GAC CTA ATA 9453 al Asp Leu Ile 3075

5	GGA TTT GTC GTT TCT GTT GTG AAA AAA ACA GGA CTT GCC CGT TTG GTG GTT GTC GTT GTG AAA AAA ACA GGA CTT GCC CGT TTG GTG GTG AAA AAA ACA GGA CTT GCC CGT TTG GTG GTG AAA AAA ACA GGA CTT GCC CGT TTG GTG GTG AAA AAA ACA GGA CTT GCC CGT TTG GTG GTG AAA AAA ACA GGA CTT GCC CGT TTG GTG TTG GTG TTG TTG TTG	9501
	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
10	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT AIG THA ATT GAT GAT AGE ASP Leu Asp Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115	9597
15	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
20	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
25	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
30	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
35	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
40	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Ty Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
45	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	9981
	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265	10029
50	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	10077
55	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
60	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys 3300 3305 3310 3315	10173

	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 10221 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	
5	TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 10269 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	
10	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 10317 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3355 3360	
15	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 10365 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	
20	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3385 3390 3395	
25	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 10461 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	
25	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	15
30	(2) INFORMATION FOR SEQ ID NO:11:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
40	<pre>(ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:</pre>	
	Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys	
45	1 5 10 15 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe	
	Thr Arg Cys Ash Lys Ala Asp 250 30 25 30 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu	
	Glu Glu Leu Ser Ser Glu Ala 110 110 171 181 45 35 40 45 Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr	
50	55 60	
	Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 75 80 65 70 75 Ser Pro Val Lys	
55	Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys 85 90 95 Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser	
	Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Ash vai 110 1101	
	100 105	
	100 105 110 Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp 120 125	
60	Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp 120 125 125 125 127 128 129 129 129 129 129 129 129 120 125	

	145					150					155					160
	145 Cys	Gly	Ser	Leu	Phe 165	150 His	Thr	Pro	Lys	Phe 170		ГÀз	Gly	Arg	Gln 175	
5	Pro	ГЛЗ	His	Ile 180		Glu	Ser	Leu	Gly 185		Glu	Val	Asp	Pro 190	Asp	Met
		_	195					200					205	Ser		
10		210					215					220		Pro		
	225					230					235			Glu		240
	-	-		_	245					250				Glu	255	
15			_	260					265					Ser 270		
			275					280					285	Ser Thr		
20		290					295					300		Lys		
	305					310					315			His		320
25					325					330				Glu	335	
	Ser	Phe	Val	340 Ser	Glu	Val	Glu	Pro	345 Asn	Asp	Thr	Авр.	Pro	350 Leu	Asp	Ser
	Asn	Val	355 Ala	His	Gln	Lys	Pro	360 Phe	Glu	Ser	Gly	Ser	365 A sp	Lys	Ile	Ser
30	Lys	370 Glu	Val	Val	Pro		375 Leu	Ala	Cys	Glu	_	380 Ser	Gln	Leu	Thr	
	385 Ser	Gly	Leu	Asn			Gln	Met	Glu		395 Ile	Pro	Leu	Leu		400 Ile
35	Ser	Ser	Сув	Asp	405 Gln		Ile	Ser	Glu 425	410 Lys	Asp	Leu	Leu	Asp 430	415 Thr	Glu
	Asn	Lys	Arg		Lys	Asp	Phe	Leu 440		Ser	Glu	Asn	Ser 445	Leu	Pro	Arg
40	Ile	Ser 450		Leu	Pro	Lys	Ser 455		Lys	Pro	Leu	Asn 460		Glu	Thr	Val
	465					470					475			Thr		480
					485					490				Val	495	
45				500					505					Glu 510		
	_		515					520					525	Asp		
50		530	_				535					540		Ile Ile		
	545	-			_	550					555			Leu		560
55	_				565					570				Phe	575	
				580					585					590 Pro		
			595	_				600				Phe	605	Ala		
60	Phe 625	610 Glu	Ala	Pro	Leu	Thr 630	615 Phe	Ala	Asn	Ala	Asp 635	620 Ser	Gly	Leu	Leu	His 640

				_	645					650				Glu	655	
5				660					665					Сув 670		
			675	_				680					685	Leu		
10	-	690		_			695					700		Ile		
10	705		_			710					715			Glu Leu		720
		-		_	725					730				Asp	735	
15		_		740					745					750 Ala		
	Leu	Ile	755 Leu	Thr	Pro	Thr		760 Lys	Asp	Val	Leu		765 Asn	Leu	Val	Met
20		770 Ser	Arg	Gly	Lys		775 Ser	Туг	Lys	Met		780 Asp	Lys	Leu	Lys	Gly 800
	785 Asn	Asn	Tyr	Glu	Ser 805	790 Asp	Val	Glu	Leu	Thr 810	795 Lys	Asn	Ile	Pro	Met 815	
25	Lys	Asn	Gln	Asp 820		Cys	Ala	Leu	Asn 825		Asn	Tyr	Lys	Asn 830		Glu
			835					840					845	Ser		
		850					855					860		Lys		
30	865					870					875			Asp		880
					885					890				Val His	895	
35				900					905					910 Thr		
	_		915	-				920					925	Ser		
40	-	930 Asp	Leu	Val	Tyr		935 Leu	Ala	Glu	Glu		940 Lys	Asn	Ser	Val	
	945 Gln	His	Ile	Lys		950 Thr	Leu	Gly	Gln	Asp 970	955 Leu	Lys	Ser	Asp	Ile 975	960 Ser
45	Leu	Asn	Ile	Asp 980	965 Lys	Ile	Pro		Lys 985	Asn	Asn	qaA		Met 990	Asn	Lys
	Trp	Ala	Gly 995		Leu	Gly	Pro		Ser		His	Ser	Phe 1005		Gly	Ser
		101)				1015	5		_		1020)	His		
50	Lys 102	-	Ser	Lys	Met	Phe 1030		Lys	Asp	Ile	Glu 1035		Gln	Tyr	Pro	Thr 104
			Ala	Cys	Val			Val	Asn	Thr			Leu	Asp	Asn	
	Lare	Larg	T.e.u	Ser	104		Gln	Ser	Tle	1050		Val	Ser	Ala	1059 His	
55	_	•		106	0				106	5				1070 Ile)	
			107	5				108	0				108	5		
		109	0			-	109	5				1100)	Asn		
60	Pro		Gln	Lys	Ala	Glu 111		Thr	Glu	Leu	Ser 111		Ile	Leu	Glu	Glu 112
			Ser	Gln	Phe			Thr	Gln	Phe			Pro	Ser	Tyr	

					1125					1130)				1135	
			Lys	1140)				1145	5				1150		
5	-		Thr 1155					1160)				1165	i		
		1170	Pro				1179	5				1180)			
10	1185	;	Glu			1190)				1199	5				120
			Ser		1205					1210)				1215	
	_	_	Phe	1220)				1225	5				1230		
15			Gln 1235					1240)				1245	5		
		1250	Thr				125	5				1260)			
20	126		Asp			1270)				1275	5				128
	_		Val		1285	,				1290)				1295	i
			Glu	1300	3				1309	5				1310	1	
25	_		Arg 1315	;				1320)				1325	5		
	_	133	Ser 0				133	5				1340)			
30	134	5	Val Asn			1350)				1359	5				136
			Asn		1365	;				1370)				1375	•
25			Ala	1386	0	•			138	5				1390)	
35			1395 Ala	5				1400	0				140	•		
		141					141	5				142	0			
40	142	5	Phe			143	0				143	5				144
			Asn		1449	5				145	0				1455)
45			Met	116	^				146	5				1470)	
43			147! Leu	5				148	0				148	5		
		149					145	_				150	0			
50	150	5	Gly			151	0				151	5				152
			Leu		152	5				153	0				153	•
55			Glu	154	0				154	5				1550)	
			155 Glu	5				156	0				156	5		
		157					157	5				158	0			Asn
60	158	15	a Asn			159	0			· Val	159 Val	5 Pro	Pro		Leu	160 Leu
	_					5					0				161	2

		_		1620)				1629	5				Ser 1630)	
5	Ile	Phe	Leu 1635		Val	Lys	Val	His 1640		Asn	Val	Glu	Lys 164	Glu 5	Thr	Ala
	Lys	Ser 1650		Ala	Thr	Сув	Tyr 1655		Asn	Gln	Ser	Pro 1660	_	Ser	Val	Ile
	Glu 166		Ser	Ala	Leu	Ala 1670		Tyr	Thr	Ser	Cys 1679		Arg	Lys	Thr	Ser 168
10	Val	Ser	Gln	Thr	Ser 1685		Leu	Glu	Ala	Lys 1690		Trp	Leu	Arg	Glu 1699	
			_	1700)				1705	5				Tyr 1710)	
15		-	1715	5				1720)				172			
	-	1730)				1735	5				1740)	Asn		
	Met 174		Asn	Ser	Tyr	Ser 1750		His	Ser	Asp	Glu 175		Tyr	Asn	Asp	Ser 176
20	•	-			1765	5	_			1770)			Pro	1775	5
	ГÀЗ	Asn	Val	Glu 1780	_	Gln	Lys		Thr 1785		Phe	Ser	Lys	Val 1790		Ser
25	Asn	Val	Lys 1799		Ala	Asn	Ala	Tyr 1800		Gln	Thr	Val	Asn 1809	Glu 5	Asp	Ile
	_	1810)				1819	5				1820)	Asn		
	182	5		_		1830)				1835	5		Glu		184
30	Pro	Pro	Ala	Phe	Arg 1849		Ala	Ser	Gly	Lys 1850		Val	Cys	Val	Ser 1855	
	Glu	Thr	Ile	Lys 1860		Val	Lys	Asp	Ile 1865		Thr	Asp	Ser	Phe 1870		Lys
35	Val	Ile	Lys 1875		Asn	Asn	Glu	Asn 1880		Ser	Lys	Ile	Cys 1885	Gln 5	Thr	Lys
	Ile	Met 1890		Gly	Cys	_	Glu 1895		Leu	Asp	Asp	Ser 1900		Asp	Ile	Leu
	190	5			_	1910	?				1915	5		His		192
40	Phe	Ala	Asp	Ile	Gln 1925		Glu	Glu	Ile	Leu 1930		His	Asn	Gln	Asn 1935	
		_		1940	,				1945	5				Val 1950)	
45														His 5		
	Val	Ser 1970		Ala	Asn	Thr	Cys 1975		Ile	Phe	Ser	Thr 1980		Ser	Gly	Lys
	Ser 198		Gln	Val	Ser	Asp 1990		Ser	Leu	Gln	Аын 1995		Arg	Gln	Val	Phe 200
50	Ser	Glu	Ile	Glu	Asp 2005		Thr	Lys	Gln	Val 2010		Ser	ГÀв	Val	Leu 2015	
	Lys	Ser	Asn	Glu 2020		Ser	Ąsp	Gln	Leu 2025		Arg	Glu	Glu	Asn 2030		Ala
55		_	203	5				2040)				204	-		
	Val	Val 205		Ser	Ser	Ala	Phe 205		Gly	Phe	Ser	Thr 2060		Ser	Gly	Lys
	206	5				207	כ				207	5	_	Gly		208
60	Glu	Glu	Phe	Asp	Leu 208		Arg	Thr	Glu	His 209		Leu	His	Tyr	Ser 2095	
	Thr	Ser	Arg	Gln	Asn	Val	Ser	Lys	Ile	Leu	Pro	Arg	Val	Asp	Lys	Arg

				210	n				210	5				211	n	
	Aan	Pro	Glu			Val	Aen	Ser			Glu	Lvg	Thr		_	Lys
	*****		211		C, S	•		212				<i>-1, -</i>	212			-1-
5	Glu	Phe		_	Ser	Asn	Asn			Val	Glu	Gly			Ser	Glu
		213	-				213					214	-			
	Asn	Asn	His	Ser	Ile	Lys	Val	Ser	Pro	Tyr	Leu	Ser	Gln	Phe	Gln	Gln
	214					215					215					216
	Asp	Lys	Gln	Gln	Leu	Val	Leu	Gly	Thr	Lys	Val	Ser	Leu	Val	Glu	Asn
10					216					217					217	
	Ile	His	Val	Leu	Gly	Lys	Glu	Gln			Pro	Lys	Asn	Val	Lys	Met
				218					218					219		
	Glu	Ile			Thr	Glu	Thr			Asp	Val	Pro			Thr	Asn
1.5			219	_	_	_,	_	220		_	_	~-	220		_,	
15	11e			Cys	Ser	Thr	-		ьys	Asp	ser			Tyr	Pne	Glu
	mb	221		17-7	~1	T1.	221		*1.	Dho	Mot	2220		7 ~~	01	T 011
	2229		ATA	val	GIU	2230		гÀв	Ald	Pne	223		Asp	Asp	GIU	Leu 224
			Ser	Lve	T.011			Hic	Δla	Thr			Len	Phe	Thr	
20	1111	no P	501	-ys	2245		501	1110	nıu	225		561	Deu	1110	225	_
	Pro	Glu	Asn	Glu			Val	Leu	Ser			Ara	Ile	Gly		
				226					226!			5		2270		5
	Arg	Gly	Glu			Ile	Leu	Val			Pro	Ser	Ile	Lys	Arq	Asn
	_	-	2275					228	_				228	_	_	
25	Leu	Leu	Asn	Glu	Phe	Asp	Arg	Ile	Ile	Glu	Asn	Gln	Glu	Lys	Ser	Leu
		229)				229	5				2300)			
	Lys	Ala	Ser	Lys	Ser	Thr	Pro	Asp	Gly	Thr	Ile	Lys	Asp	Arg	Arg	Leu
	2305	-				2310					2315					232
	Phe	Met	His	His			Leu	Glu	Pro			Cys	Val	Pro		_
30	_,		_		2325			_,		2330		_		_,	2335	
	Thr	Thr	гля			GIn	GIU	11e			Pro	Asn	Pne	Thr		Pro
	a 1	~1~	~1	2340		C	T	00**	2345		TT	~1	TT	2350 Leu		Lon
	GIY	GIII	2355		Leu	Sei	_	2360		Leu	IÀT	Gru	2365		TIIL	Lea
35	Glu	Lvs			Ser	Agn				Ser	Glv	His		Phe	Tvr	Gln
		2370					2375				,	2380			-1-	
	Val	Ser	Ala	Thr	Arq	Asn			Met	Arq	His			Thr	Thr	Gly
	2385				_	2390		•		_	2395					240
	Arg	Pro	Thr	Lys	Val	Phe	Val	Pro	Pro	Phe	Lys	Thr	Lys	Ser	His	Phe
40					2405					2410					2415	
	His	Arg	Val			Сув	Val	Arg			Asn	Leu	Glu	Glu		Arg
	_		_	2420			_		2425					2430		_
	Gln	Lys												Lys		Lys
4 5	77 -	3														01 -
45	TTE	Asn 245(_	Asn	GIU	TIE	H18		Pne	Asn	rÀs	2460		ser	ASI	Gln
	Δla			V=1	Thr	Dha			Cva	G) 11	Glu			Leu	Δen	T.en
	2465		VIG	Vai	TILL	2470		цуз	Cys	Giu	2475		FIO	neu	чар	248
			Ser	Leu	Gln			Ara	Asp	Ile			Met	Arg	Ile	
50					2485					2490				5	2499	
	Lys	Lys	Gln	Arg			Val	Phe	Pro			Gly	Ser	Leu		
	-	-		2500)	_			2505	5		_		2510)	
	Ala	Lys	Thr	Ser	Thr	Leu	Pro	Arg	Ile	Ser	Leu	Lys	Ala	Ala	Val	Gly
			2515					2520					2525			
55	Gly			Pro	Ser	Ala	-		His	Lys	Gln		_	Thr	Tyr	Gly
		2530			_	_	2535		_	_	_	2540			_	
			Lys	His	Cys			Ile	Asn	Ser			Ala	Glu	ser	
	2545		u:-	mь	01	2550		nh -	03 -	7	2555		T	т~~	かん~	256
60	GIU	rne	uta	inr	2565	_	TAL	rne	GTA	Lуs 2570		ser	ьец	Trp	2575	
00	Live	GIV	Tle	Gln	_		Agn	Glv	Glv			Tlo	Pro	Ser		
	-,5	1	-10	2580			P	1	258	_		-10		2590		p
					-					-						

	GIY	ràs	2599		гув	GIU	GIU	2600	-	Arg	AIA	Leu	260!	_	Thr	PIC
5		2610)			Leu	261	5				262	0			
	Arg 2625		Ile	Ile	Trp	Lys 263		Ala	Ala	Met	Glu 263		Ala	Phe	Pro	Lys 264
	Glu	Phe	Ala	Asn	Arg 2649	Cys 5	Leu	Ser	Pro	Glu 265	-	Val	Leu	Leu	Gln 265	
10	Lys	Tyr	Arg	Tyr 2660		Thr	Glu	Ile	Asp 2669		Ser	Arg	Arg	Ser 267		Ile
	Lys	Lys	1le 2679		Glu	Arg	Asp	Asp 2680		Ala	Ala	ГÀа	Thr 2689		Val	Leu
15	Сув	Val 2690		Asp	Ile	Ile	Ser 269		Ser	Ala	Asn	Ile 270		Glu	Thr	Ser
	Ser 2709		Lys	Thr	Ser	Ser 271		Asp	Thr		Lys 271		Ala	Ile	Ile	Glu 272
	Leu	Thr	Asp	Gly	Trp 2729	Tyr	Ala	Val	Lys	Ala 2730		Leu	Asp	Pro	Pro 273	
20	Leu	Ala	Val	Leu 2740	_	Asn	Gly	Arg	Leu 2749		Val	Gly	Gln	Lys 275		Ile
	Leu	His	Gly 2755		Glu	Leu	Val	Gly 2760		Pro	Asp	Ala	Cys 2769		Pro	Leu
25	Glu	Ala 2770		Glu	Ser	Leu	Met 277		Lys	Ile	Ser	Ala 2780		Ser	Thr	Arg
	Pro 2785		Arg	Trp	Tyr	Thr 2790		Leu	Gly	Phe	Phe 2795		Asp	Pro	Arg	Pro 280
	Phe	Pro	Leu	Pro	Leu 2809	Ser	Ser	Leu	Phe	Ser 2810		Gly	Gly	Asn	Val 2815	
30	Cys	Val	Asp	Val 2820		Ile	Gln	Arg	Ala 2825	-	Pro	Ile	Gln	Trp 2830		Glu
	Lys	Thr	Ser 2835		Gly	Leu	Tyr	Ile 2840		Arg	Asn	Glu	Arg 2845		Glu	Glu
35	Lys	Glu 2850		Ala	Lys	Tyr	Val 2855		Ala	Gln	Gln	Lys 2860		Leu	Glu	Ala
	Leu 2869		Thr	Lys	Ile	Gln 2870		Glu	Phe	Glu	Glu 2879		Glu	Glu	Asn	Thr 288
					2885					2890)				2895	5
40	Ala	Leu	Gln	Asp 2900	_	Ala	Glu	Leu	Tyr 2909		Ala	Val	ГХE	Asn 2910		Ala
	Asp	Pro	Ala 2915	-	Leu	Glu	Gly	Tyr 2920		Ser	Glu	Glu	Gln 2925		Arg	Ala
45	Leu	Asn 2930		His	Arg	Gln	Met 2935		Asn	Asp	Lys	Lys 2940		Ala	Gln	Ile
	Gln 2945		Glu	Ile	Arg	Lys 2950		Met	Glu	Ser	Ala 2955		Gln	ГÀв	Glu	Gln 296
	Gly	Leu	Ser	Arg	Asp 2965	Val 5	Thr	Thr	Val	Trp 2970		Leu	Arg	Ile	Val 2975	
50	Tyr	Ser	Lys	Lys 2980		Lys	Asp	Ser	Val 2985		Leu	Ser	Ile	Trp 2990		Pro
	Ser	Ser	Asp 2995		Tyr	Ser	Leu	Leu 3000		Glu	Gly	Lys	Arg 3005	_	Arg	Ile
55	Tyr	His 3010		Ala	Thr	Ser	Lys 3015		Lys	Ser	Lys	Ser 3020		Arg	Ala	Asn
	Ile 3029		Leu	Ala	Ala	Thr 3030	_	ГÀа	Thr	Gln	Tyr 3035		Gln	Leu	Pro	Val 304
	Ser	Asp	Glu	Ile	Leu 3045	Phe		Ile	Tyr	Gln 3050	Pro		Glu	Pro	Leu 3055	His
60	Phe	Ser	Lys	Phe 3060	Leu	Asp	Pro	Asp	Phe 3065	Gln		Ser	Сла	Ser 3070	Glu	
	Asp	Leu	Ile			Val	Val	Ser			Lys	Lys	Thr			Ala

		3075	i			3080)				3085	5		
	Pro Phe)	-		309	5				3100)			
5	Phe Trp 3105		-	3110	0				3119	5				312
	Ala Ala		31	L25				3130)				3135	5
10	Thr Leu		3140				3145	5				3150)	
	Gly His	3155	,			3160)				3169	5		
	Ile Asp	ס			317	5				3180)			
15	His Ala 3185			3190	0				3195	5				320
	Gly Pro		32	205				3210)				3215	5
20	Met Ser		3220				3229	5				3230)	
	Cys Met	3235	;			3240)				3245	5		
	Thr Ser	0			325	5				3260)			
25	Cys Lys 3265			3270	0				3275	5				328
	Pro Pro		32	285				3290)				3295	5
30	Ala Phe		3300				3309	5				3310)	
	Lys Lys	3315	i			3320)				3325	5		
	Asn Glu 333	0			333	5				3340)			
35	Ala Leu 3345			3350	0				3355	5				336
	Gln Phe		3 3	365				3370)				3375	5
40	Glu Asp	-	3380				3389	5				3390)	
	Glu Gln	3395	;			3400)		Glu	Cys	Glu 3405		Asn	Lys
	Gln Asp		Ile Th	ir Thr	Lys 341		Tyr	Ile						
45		(2)	INFO	RMATIO	N FOI	R SE	Q ID	NO:	12:					
	(.	i) SE	QUENCI	CHAR	ACTE	RIST:	cs:							
50		(B)	TYPE:	H: 104	ic a	cid		3						
				DEDNES: DGY: 1			2							
				LE TYP	E: c	AND								
55	(•	FEATUR		a - 7º	-								
		(B)	LOCA'	/KEY:	229.	10	482							
60				R INFO										
	(xi) S	SEQUEN	CE DES	CRIP	TION	: SE	Q ID	NO:	12:				

5	GGTGGCG TCTGCTG ACAGATT CTGGAGG	TGT (CTCG(GACC(GGTG:	CG G	TTTG(TTTT	CGGC TGTC	g gt a gc'	GGGT TTAC	CGCC TCCG	GCC	ggga Aaaa	GAA AAG . A AT	GCGT AACT G CC	GAGGGG GCACCT	60 120 180 237
10	GGA TCC Gly Ser 5															285
15	AAC AAA Asn Lys 20															333
2.0	TCT TCA Ser Ser															381
20	CAT AAA His Lys															429
25	AAA CCA Lys Pro															477
30	CAA GGG Gln Gly 85															525
35	AAA TTC Lys Phe 100															573
40	AGT CTT Ser Leu									_	_			_		621
40	TGT CCA															669
45	TGT ACA															717
50	TTG TTT Leu Phe 165	His														765
55	ATT TCT Ile Ser 180															813
60	AGT TCT Ser Ser													_	_	861
60	AGA AAT Arg Asn															909

		215				220			225			
5		AGC Ser										957
10		ATC Ile										1005
15		AGT Ser										1053
		TGC Cys										1101
20		GTA Val 295										1149
25		TGT Cys										1197
30		AAG Lys										1245
35		AAA` Lys	Ser									1293
		GAA Glu										1341
40		CCC Pro 375		_	_			_		_	_	1389
45		TTG Leu										1437
50		CAG Gln										1465
55		ATT Ile										1533
33		TTT Phe										1581
60		TCG Ser 455										1629

5													TGC Cys 480				1677
10													TCT Ser				1725
10	GGT Gly 500	ATC Ile	AAA Lys	AAG Lys	TCT Ser	ATA Ile 505	TTC Phe	AGA Arg	ATA Ile	AGA Arg	GAA Glu 510	TCA Ser	CCT Pro	AAA Lys	GAG Glu	ACT Thr 515	1773
15													AAC Asn				1821
20													ACT Thr				1869
25													AAT Asn 560				1917
30													AAT Asn				1965
30													TAT Tyr				2013
35													GAC Asp				2061
40													GCT Ala				2109
45													CAT His 640				2157
50													ACT Thr				2205
													AGA Arg				2253
55													TAT Tyr				2301
60													CCA Pro				2349

	TCT Ser	CTG Leu	TCA Ser 710	TGC Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	CAG Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
5	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA AJ.a	GCA Ala	TGT Cys	CAC His	2445
10	CCA Pro 740	GTA Val	CAA Gln	CAC His	TCA Ser	AAA Lys 745	GTG Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	ACT Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
15	CAG Gln	AAA Lys	AGT Ser	CTT Leu	TTA Leu 760	TAT Tyr	GAT Asp	CAT His	GAA Glu	AAT Asn 765	GCC Ala	AGC Ser	ACT Thr	CTT Leu	ATT Ile 770	TTA Leu	2541
20	Thr	Pro	Thr	Ser 775	Lys	Asp	Val	Leu	Ser 780	Asn	Leu	Val	ATG Met	785	ser	Arg	2589
25	Gly	Lys	Glu 790	Ser	Tyr	Lys	Met	Ser 795	Asp	Lys	Leu	Lys	GGT Gly 800	Asn	Asn	ıyr	2637
	Glu	Ser 805	Asp	Val	Glu	Leu	Thr 810	Lys	Asn	Ile	Pro	Met 815	GAA Glu	гÀг	Asn	GIN	2685
30	Asp 820	Val	Сув	Ala	Leu	Asn 825	Glu	Asn	Tyr	Lys	Asn 830	Val	GAG Glu	Leu	Leu	835	2733
35	Pro	Glu	Lys	Tyr	Met 840	Arg	Val	Ala	Ser	Pro 845	Ser	Arg	AAG Lys	vaı	850	PILE	2781
40	Asn	Gln	Asn	Thr 855	Asn	Leu	Arg	Val	Ile 860	Gln	Lys	Asn	CAA Gln	865	GIU	IIII	2829
45	Thr	Ser	1le 870	Ser	Lys	Ile	Thr	Val 875	Asn	Pro	Asp	ser	GAA Glu 880	GIU	Leu	PHE	2877
	Ser	Asp 885	Asn	Glu	Asn	Asn	Phe 890	Val	Phe	Gln	Ile	Ala 895		GIu	Arg	Asn	2925
50	Asn 900	Let	ı Ala	Lev	Gly	905	Thr	гуs	Glu	Leu	910	Glu	ACA Thr	Asp	Leu	915	2973
55	Сує	val	l Asr	ı Glu	920)	Phe	: Гуз	as as a	925	Thr	Met	GTT Val	Leu	930	GIY	3021
60	Ası	Th:	r Gly	93!	D Lys	s Glr	n Ala	a Thi	940	val	. Ser	: Ile	. гЛа	945	Asp	Leu	3069
	GT'	r TA	T GT	r cr	r gc	A GAG	G GAG	AA E	C AA	AA.	r AGT	GTA	AAG	CAG	CAT	ATA '	3117

	Val	туг	Val 950	Leu	Ala	Glu	Glu	Asn 955	Lys	Asn	Ser	Val	Lys 960	Gln	His	Ile	
5	AAA Lys	ATG Met 965	ACT Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	TTA Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	TCC Ser	TTG Leu	TAA NaA	ATA Ile	3165
10	GAT Asp 980	Lys	ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	TAA naA	AAT Asn	GAT Asp	TAC Tyr	ATG Met 990	GAC Asp	AAA Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
15	CTC Leu	TTA Leu	GGT Gly	Pro	ATT Ile L000	TCA Ser	AAT Asn	CAC His	Ser	TTT Phe 1005	GGA Gly	GGT Gly	AGC Ser	Phe	AGA Arg 1010	ACA Thr	3261
	GCT Ala	TCA Ser	AAT Asn	AAG Lys 1015	GAA Glu	ATC Ile	AAG Lys	Leu	TCT Ser 1020	GAA Glu	CAT His	AAC Asn	Ile	AAG Lys 1025	AAG Lys	AGC Ser	3309
20	AAA Lys	Met	TTC Phe 1030	TTC Phe	AAA Lys	GAT Asp	Ile	GAA Glu 1035	GAA Glu	CAA Gln	TAT Tyr	Pro	ACT Thr 1040	AGT Ser	TTA Leu	GCT Ala	3357
25	TGT Cys	GTT Val	GAA Glu	ATT Ile	GTA Val	Asn	ACC Thr 1050	TTG Leu	GCA Ala	TTA Leu	Asp	AAT Asn 1055	CAA Gln	AAG Lys	AAA Lys	CTG Leu	3405
30	AGC Ser 1060	Lys	CCT Pro	CAG Gln	Ser	ATT Ile 1065	Asn	ACT Thr	GTA Val	Ser	GCA Ala 1070	CAT His	TTA Leu	CAG Gln	Ser	AGT Ser 1075	3453
35	GTA Val	GTI Val	GTT Val	Ser	GAT Asp 1080	Cys	AAA Lys	AAT Asn	Ser	CAT His 1085	ATA Ile	ACC Thr	CCT Pro	Gln	ATG Met 1090	TTA Leu	3501
	TTT Phe	TCC Ser	AAG Lys	CAG Gln 1095	Asp	TTT Phe	AAT Asn	Ser	AAC Asn 1100	CAT	AAT Asn	TTA Leu	Thr	CCT Pro 1105	AGC Ser	CAA Gln	3549
40	AAC Lys	G GCA	A GAA A Glu 1110	Ile	ACA Thr	GAA Glu	Leu	TCT Ser 1115	Thr	ATA Ile	TTA Leu	Glu	GAA Glu 1120	TCA Ser	GGA Gly	AGT Ser	3597
45	CA(3 TTT n Phe 1125	e Glu	TTT Phe	ACT Thr	CAG	TTT Phe	Arg	AAA Lys	CCA Pro	Ser	TAC Tyr 1135	Ile	TTG Leu	CAG Gln	AAG Lys	3645
50	AG: Se:	r Th	A TTT	GAA	GTG Val	CCT Pro	Glu	AAC Asn	CAG Gln	ATG Met	ACT Thr	Ile	TTA Leu	AAG Lys	ACC	ACT Thr 1155	3693
55	TC' Se:	T GAG	G GAZ u Glu	TGC Cys	AGA Arg 1160	, Asp	GCI Ala	GAT Asp	CTT Leu	CAT His 1165	. Val	ATA Ile	ATG Met	AAT Asn	GCC Ala 1170	CCA Pro	3741
	TC Se	G AT	T GGT e Gly	r CAC y Gli 1175	ı Val	A GAG L As <u>r</u>	C AGO Sei	AGC Sei	AAG Lys 1180	Glr	TTI Phe	GAZ Glu	GGT Gly	ACA Thr	. Val	GAA Glu	3789
60	AT Il	T AA e Ly	A CGG	G AAG g Ly	G TT	r GCT e Ala	r GGC a Gly	CTC	TTC	AA Lys	IAA A	GA(TGT Cys	AAC aas	AAA 1 Lys	AGT Ser	3837

PCT/US98/16905 WO 99/09164

GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Jos Thr Glu Ala Leu Gln AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala ACT ARA ACG GAG CAA AAT ATA ARA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT

Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe

5	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAT TTO GTA AAT AAA AAT AAT AAA AAT AAT AAA AAT AA	4605
	TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475	4653
10	GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490	4701
15	AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505	4749
20	GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520	4797
25	TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535	4845
	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1545 1550 1555	4893
30	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570	4941
35	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
40	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
45	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu In Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1635	5133
50	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
55	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
60	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277

		-
	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
5	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710	5373
10	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
15	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
20	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
25	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
30	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
35	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
40	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
45	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
50	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
55	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
60	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG	6045

Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val

PCT/US98/16905 WO 99/09164

TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT

Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val

5	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
10	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala Ala 2455 2460 2465	7629
15	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu lle Thr Ser 2470 2475 2480	7677
20	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CCA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
25	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515	7773
	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
30	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545	7869
35	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560	7917
40	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
45	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595	8013
	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610	8061
50	CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625	8109
55	ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640	8157
60	AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655	8205

	TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675	8253
5	ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 2685 2690	8301
10	GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705	8349
15	ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720	8397
20	GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro ro Leu Leu Ala Val 2725 2730 2735	8445
25	TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755	8493
25	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
30	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
35	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800	8637
40	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
4.5	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val lle lle Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 2825 2830 2835	8733
45	TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala 2840 2845 2850	8781
50	GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 2860 2865	8829
55	AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880	8877
60	TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895	8925
	GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT	8973

		Tyr Glu Ala Val	Lys Asn Ala Ala Asp Pro Ala 2910 2915
5	TAC CTT GAG GGT TAT Tyr Leu Glu Gly Tyr 2920	Phe Ser Glu Glu	CAG TTA AGA GCC TTG AAT AAT 9021 Gln Leu Arg Ala Leu Asn Asn 925 2930
10	CAC AGG CAA ATG TTG	AAT GAT AAG AAA	CAA GCT CAG ATC CAG TTG GAA 9069
	His Arg Gln Met Leu	Asn Asp Lys Lys	Gln Ala Gln Ile Gln Leu Glu
	2935	2940	2945
15	ATT AGG AAG ACC ATG	GAA TCT GCT GAA	CAA AAG GAA CAA GGT TTA TCA 9117
	Ile Arg Lys Thr Met	Glu Ser Ala Glu	Gln Lys Glu Gln Gly Leu Ser
	2950	2955	2960
	AGG GAT GTC ACA ACC	GTG TGG AAG TTG	CGT ATT GTA AGC TAT TCA AAA 9165
	Arg Asp Val Thr Thr	Val Trp Lys Leu	Arg Ile Val Ser Tyr Ser Lys
	2965	2970	2975
20	Lys Glu Lys Asp Ser	GTT ATA CTG AGT Val Ile Leu Ser 2985	ATT TGG CGT CCA TCA TCA GAT 9213 Ile Trp Arg Pro Ser Ser Asp 2990 2995
25	TTA TAT TCT CTG TTA Leu Tyr Ser Leu Leu 3000	Thr Glu Gly Lys	AGA TAC AGA ATT TAT CAT CTT 9261 Arg Tyr Arg Ile Tyr His Leu 3005 3010
30	GCA ACT TCA AAA TCT	AAA AGT AAA TCT	GAA AGA GCT AAC ATA CAG TTA 9309
	Ala Thr Ser Lys Ser	Lys Ser Lys Ser	Glu Arg Ala Asn Ile Gln Leu
	3015	3020	3025
35	GCA GCG ACA AAA AAA	ACT CAG TAT CAA	CAA CTA CCG GTT TCA GAT GAA 9357
	Ala Ala Thr Lys Lys	Thr Gln Tyr Gln	Gln Leu Pro Val Ser Asp Glu
	3030	3035	3040
	ATT TTA TTT CAG ATT	TAC CAG CCA CGG	GAG CCC CTT CAC TTC AGC AAA 9405
	Ile Leu Phe Gln Ile	Tyr Gln Pro Arg	Glu Pro Leu His Phe Ser Lys
	3045	3050	3055
40	Phe Leu Asp Pro Asp	TTT CAG CCA TCT Phe Gln Pro Ser 3065	TGT TCT GAG GTG GAC CTA ATA 9453 Cys Ser Glu Val Asp Leu Ile 3070 3075
45	GGA TTT GTC GTT TCT Gly Phe Val Val Ser 3080	Val Val Lys Lys	ACA GGA CTT GCC CCT TTC GTC 9501 Thr Gly Leu Ala Pro Phe Val 3085 3090
50	TAT TTG TCA GAC GAA	TGT TAC AAT TTA	CTG GCA ATA AAG TTT TGG ATA 9549
	Tyr Leu Ser Asp Glu	Cys Tyr Asn Leu	Leu Ala Ile Lys Phe Trp Ile
	3095	3100	3105
55	GAC CTT AAT GAG GAC	ATT ATT AAG CCT	CAT ATG TTA ATT GCT GCA AGC 9597
	Asp Leu Asn Glu Asp	o Ile Ile Lys Pro	His Met Leu Ile Ala Ala Ser
	3110	3115	3120
	AAC CTC CAG TGG CGA	A CCA GAA TCC AAA	A TCA GGC CTT CTT ACT TTA TTT 9645
	Asn Leu Gln Trp Arg	g Pro Glu Ser Lys	S Ser Gly Leu Leu Thr Leu Phe
	3125	3130	3135
60	GCT GGA GAT TTT TC	r GTG TTT TCT GCT	F AGT CCA AAA GAG GGC CAC TTT 9693
	Ala Gly Asp Phe Se	r Val Phe Ser Ala	A Ser Pro Lys Glu Gly His Phe

	3140				:	3145				:	3150				;	3155	
5		GAG Glu		Phe					Asn					Ile			9741
10		TGC . Cys	Asn					Lys					Leu				9789
15	_	CCC Pro					Pro					Thr					9837
15	Thr	GCT Ala 3205				Pro					Lys						9885
20		AAT Asn			Ile					Pro					Met		9933
25		AGG .		Ser					Val					Thr			9981
30		TGT :	Lys					Ile					Asn				10029
35		AGA (Leu					Leu					10077
33	Ser	CCC Pro				Phe					Ala						10125
40		CCA ?			Cys					Glu					Lys		10173
45	GAA Glu	CTG :	AAT Asn	Ser	CCT Pro 3320	CAG Gln	ATG Met	ACT Thr	Pro	TTT Phe 325	AAA Lys	AAA Lys	TTC Phe	Asn	GAA Glu 3330	ATT Ile	10221
50		CTT '	Leu					Ile			-		Leu				10269
		ACC Thr					Ser					Glu					10317
55	Ser	GTC : Val :				Thr					Thr						10365
60		AGA Arg			Arg					Ser					Gln		10413

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 5 3400 3405 10485 ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415 10 (2) INFORMATION FOR SEQ ID NO:13: (i) SEQUENCE CHARACTERISTICS: 15 (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 20 (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13: Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 25 10 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 30 40 Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 55 Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 35 Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys 90 Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser 105 100 Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp 40 120 125 Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val 135 140 Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val 150 Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr 45 165 170 Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met 185 Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val 50 200 Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp 215 220 Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu 230 235 55 Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr 245 250 Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn 265 Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro 60 275 280 His Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu

300

295

	305					310					315					Leu 320
5		_		_	Thr 325					330					335	
				340	Сув				345					350		
1.0			355		Glu			360					365			
10		370			Gln	_	375					380				Leu
	385				Gly	390					395					400
15		_			405 Gln					410					415	
	Asn	Lys	Arg	420 Lys	Lys	Asp	Phe		425 Thr	Ser	Glu	Asn		430 Leu	Pro	Arg
20	Ile		435 Ser	Leu	Pro	Lys		440 Glu	ГÀа	Pro	Leu	Asn 460	445 Glu	Glu	Thr	Val
	Val	450 Asn	Lys	Arg	Asp	Glu 470	455 Glu	Gln	His	Leu	Glu 475		His	Thr	Asp	Cys 480
25		Leu	Ala	Val	Lys 485		Ala	Ile	Ser	Gly 490		Ser	Pro	Val	Ala 495	
				500	Ile				505					510		
2.0	_		515		Asn			520					525			
30		530	_		Thr		535					540				
	545	-			Lys Ala	550	_				555					560
35	_				565 Ser					570					575	
		_		580	Glu			Tyr	585				Ile	590		
40	Gln	_	595 Ser	Glu	Leu	Ile		600 Cys	Ser	Ala	Gln		605 Glu	Ala	Asn	Ala
		610 Glu	Ala	Pro	Leu	Thr 630	615 Phe	Ala	Asn	Ala	Asp 635	620 Ser	Gly	Leu	Leu	His 640
45	625 Ser	Ser	Val	Lys	Arg 645		Cys	Ser		Asn 650	qaA	Ser	Glu		Pro 655	Thr
				660	Ser				665					670		Arg
			675	-				680					685			Tyr
50	-	690		-	Суз		695					700				
	705		_		Leu Lys	710					715					720
55					725 Val					730					735	
				740	Lys				745					750		
60			755		Pro			760				Ser	765			
	Ile	770 Ser	Arg	Gly	Lys	Glu	775 Ser	Tyr	Lys	Met	Ser	780 Asp	Lys	Leu	Lys	Gly

	785					790					795					800
		Asn	Tyr	Glu	Ser 805		Val	Glu	Leu	Thr 810		Asn	Ile	Pro	Met 815	
5	_			820		_			825				_	Asn 830		
	Leu	Leu	Pro 835	Pro	Glu	Lys	Tyr	Met 840	Arg	Val	Ala	Ser	Pro 845	Ser	Arg	Lys
10		850					855			_		860		ГÀЗ		
	865					870		_			875			Asp		880
1.5					885					890				Ile	895	
15				900					905					His 910 Thr		
	-		915	_				920					925	Ser		
20		930					935					940		Ser		
	945	_			_	950					955			Asp		960
25				•	965			_		970		_		Met	975	
				980					985					990 Gly		
			995					1000)				1005			
30		1010)				1015	•				1020)	Tyr		
	1029 Ser		Ala	Cys	Val	1030 Glu		Val	Asn	Thr	1035 Leu		Leu	Asp	Asn	104 Gln
2.5		_	_	_	1045		a1	0	T1 -	1050		17-1	Cor	77.	1055	
35	Lys	LVR	Leu	ser	TAVE		(3170					vaı	ser			
	_	_		1060)				1065	5				1070		
	Gln	Ser	Ser 1075	1060 Val) Val	Val	Ser	Asp 1080	1065 Cys)	Lys	Asn	Ser	His 1085	1070 Ile 5	Thr	Pro
40	Gln	Ser Met	Ser 1075 Leu	1060 Val Phe	Val Ser	Val Lys	Ser Gln 1099	Asp 108(Asp	1065 Cys) Phe	Lys Asn	Asn Ser	Ser Asn 1100	His 1089 His	1070 Ile S Asn	Thr Leu	Pro Thr
40	Gln Gln Pro	Ser Met 1090 Ser	Ser 1075 Leu	1060 Val Phe	Val Ser	Val Lys Glu	Ser Gln 1099	Asp 108(Asp	1065 Cys) Phe	Lys Asn	Asn Ser Ser	Ser Asn 1100 Thr	His 1089 His	1070 Ile 5	Thr Leu	Pro Thr
40	Gln Gln Pro 110 Ser	Met 1090 Ser 5	Ser 1075 Leu) Gln Ser	1060 Val Phe Lys Gln	Val Ser Ala Phe	Val Lys Glu 1110 Glu	Ser Gln 1099 Ile Phe	Asp 1080 Asp Thr	1065 Cys) Phe Glu Gln	Lys Asn Leu Phe	Asn Ser Ser 1115 Arg	Ser Asn 1100 Thr Lys	His 1085 His) Ile Pro	1070 Ile S Asn Leu Ser	Thr Leu Glu Tyr 1135	Thr Glu 112 Ile
40	Gln Gln Pro 110 Ser Leu	Ser Met 1090 Ser 5 Gly	Ser 1075 Leu Gln Ser	Phe Lys Gln Ser	Val Ser Ala Phe 1125 Thr	Val Lys Glu 1110 Glu Phe	Ser Gln 1099 Ile Phe Glu	Asp 1080 Asp Thr Thr	1065 Cys Phe Glu Gln Pro	Lys Asn Leu Phe 1130 Glu	Asn Ser Ser 1115 Arg Asn	Asn 1100 Thr Lys	His 1085 His Ile Pro Met	1070 Ile Asn Leu Ser Thr 1150	Thr Leu Glu Tyr 1135 Ile	Pro Thr Glu 112 Ile Leu
	Gln Gln Pro 110 Ser Leu Lys	Ser Met 1090 Ser Gly Gln Thr	Ser 1075 Leu Gln Ser Lys Thr	1060 Val Phe Lys Gln Ser 1140 Ser	Val Ser Ala Phe 1125 Thr	Val Lys Glu 1110 Glu Phe	Gln 1099 Ile Phe Glu Cys	Asp 1080 Asp Thr Thr Val Arg 1160	1065 Cys Phe Glu Gln Pro 1145 Asp	Lys Asn Leu Phe 1130 Glu Ala	Asn Ser Ser 1115 Arg Asn Asn	Asn 1100 Thr Lys Gln Leu	His 1085 His Ile Pro Met His 1165	1070 Ile Asn Leu Ser Thr 1150 Val	Thr Leu Glu Tyr 1135 Ile	Pro Thr Glu 112 Ile Leu Met
	Gln Gln Pro 110 Ser Leu Lys Asn	Ser Met 1090 Ser 5 Gly Gln Thr Ala 117	Ser 1075 Leu Gln Ser Lys Thr 1155 Pro	Phe Lys Gln Ser 1146 Ser Ser	Val Ser Ala Phe 1129 Thr Glu	Val Lys Glu 1110 Glu Phe Glu	Gln 1099 Ile Phe Glu Cys Gln 117	Asp 1080 Asp Thr Thr Val Arg 1160 Val	Office of the second se	Lys Asn Leu Phe 1130 Glu Ala	Asn Ser Ser 1115 Arg Asn Asn	Asn 1100 Thr Lys Gln Leu Lys 1180	His 1085 His Ile Pro Met His 1165 Gln	1070 Ile Asn Leu Ser Thr 1150 Val Phe	Thr Leu Glu Tyr 1135 Ile Ile Glu	Thr Glu 112 Ile Leu Met Gly
45	Gln Gln Pro 110 Ser Leu Lys Asn Thr	Ser Met 1090 Ser Gly Gln Thr Ala 1170 Val	Ser 1075 Leu Gln Ser Lys Thr 1155 Pro	Phe Lys Gln Ser 1146 Ser Ser	Val Ser Ala Phe 1129 Thr Glu	Val Lys Glu 1110 Glu Phe Glu	Gln 1099 11e Phe Glu Cys Gln 1179 Lys	Asp 1080 Asp Thr Thr Val Arg 1160 Val	Office of the second se	Lys Asn Leu Phe 1130 Glu Ala	Asn Ser Ser 1115 Arg Asn Asn	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu	His 1085 His Ile Pro Met His 1165 Gln	1070 Ile Asn Leu Ser Thr 1150 Val	Thr Leu Glu Tyr 1135 Ile Ile Glu	Thr Glu 112 Ile Leu Met Gly
45	Gln Gln Pro 110 Ser Leu Lys Asn Thr 118 Asn	Met 1090 Ser 5 Gly Gln Thr Ala 1170 Val 5	Ser 1075 Leu Gln Ser Lys Thr 1155 Pro Glu Ser	1060 Val Phe Lys Gln Ser 1140 Ser Ile Ala	Val Ser Ala Phe 1129 Thr Glu Ile Lys Ser	Val Lys Glu 1110 Glu Phe Glu Gly Arg 1190 Gly	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys	Asp 1080 Asp Thr Thr Val Arg 1160 Val Phe	Office of the second se	Lys Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn	His 1085 His Ile Pro Met His 1165 Gln Lys	In the second of	Thr Leu Glu Tyr 1135 Ile Glu Asp Gly 1215	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe
45	Gln Gln Pro 110 Ser Leu Lys Asn Thr 118 Asn	Met 1090 Ser 5 Gly Gln Thr Ala 1170 Val 5 Lys	Ser 1075 Leu Gln Ser Lys Thr 1155 Pro Glu Ser	1060 Val Phe Lys Gln Ser 1140 Ser Ile Ala	Val Ser Ala Phe 1129 Thr Glu Ile Lys Ser 1209 Ser	Val Lys Glu 1110 Glu Phe Glu Gly Arg 119 Gly 5	Gln 1099 Ile Phe Glu Cys Gln 117 Lys Tyr His	Asp 1080 Asp Thr Thr Val Arg 1160 Val Phe Leu	Phe Glu Gln Pro 1145 Asp Asp Ala Thr	Lys Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys	Asn Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn	His 1085 His Ile Pro Met His 1165 Gln Lys Glu Val	In the ser ser ser ser ser ser ser ser ser se	Thr Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe
4 5	Gln Gln Pro 110 Ser Leu Lys Asn Thr 118 Asn Arg	Met 1090 Ser 5 Gly Gln Thr Ala 1170 Val 5 Lys Gly Leu	Ser 1075 Leu Gln Ser Lys Thr 1155 Pro Glu Ser Phe Gln 123	1060 Val Phe Lys Gln Ser 1140 Ser Ile Ala Tyr 1220 Lys	Val Ser Ala Phe 1129 Thr Glu Ile Lys Ser 1209 Ser	Lys Glu 1110 Glu Phe Glu Gly Arg 119 Gly Arg Arg Arg Arg Val	Gln 1099 Ile Phe Glu Cys Gln 117 Lys Tyr His	Asp 1080 Asp Thr Thr Val Arg 1160 Val Fhe Leu Gly Leu 124	Office of the control	Lys Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys Ser	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn Asn	His 1085 His Ile Pro Met His 1165 Gln Lys Glu Val	In the ser that th	Thr Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser
4 5	Gln Gln Pro 110 Ser Leu Lys Asn Thr 118 Asn Arg Ala Glu	Ser Met 1090 Ser Gly Gln Thr Ala 117 Val Lys Gly Leu Glu 125	Ser 1075 Leu Gln Ser Lys Thr 1155 Pro Glu Ser Phe Gln 123 Thr	1060 Val Phe Lys Gln Ser 1140 Ser Ile Ala Tyr 1220 Lys Ser	Val Ser Ala Phe 1129 Thr Glu Ile Lys Ser 1209 Ser Ala Ala	Val Lys Glu 1110 Glu Phe Glu Gly Arg 119 Gly Ala Val	Ser Gln 1099 Ile Phe Glu Cys Gln 117 Lys Tyr His Lys Val 125	Asp 1080 Asp Thr Thr Val Arg 1160 Val Fhe Leu Gly Leu 124 His	Open 1065 Cys Phe Glu Gln Pro 1145 Asp Asp Ala Thr Thr 1225 Phe Pro	Lys Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys Ser Ile	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp	Asn 1100 Thr Lys Gln Leu Lys 1180 Leu Asn Asn Ile Leu 1260	His 1085 His Ile Pro Met His 1165 Gln Lys Glu Val Glu 1245 Ser	In the ser ser ser ser ser ser ser ser ser se	Thr Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser Lys

	Lys Thi	val	Ser	Glu 1285		Asn	Asn	Lys	Cys 1290		Leu	Ile	Leu	Gln 1299	
5	Asn Ile	Glu	Met 1300		Thr	Gly	Thr	Phe 1305		Glu	Glu	Ile	Thr 1310		Asn
-	Tyr Lys	Arg 131	Asn		Glu	Asn	Glu 1320	Asp		Lys	Tyr	Thr 1329	Ala		Ser
	Arg Asi		His	Asn	Leu	Glu 1335		Asp	Gly	Ser	Asp 1340		Ser	Lys	Asn
10	Asp Thi		-		1350)	_			1355	5				136
	Gln His			1365	5				1370)				1375	5
15	Asn Thi		1380)				1385	5				1390)	
	Ala Lys	139	5				1400)				1405	5		
	Leu Thi	LO		_		1415	5				1420)			
20	Asp Thi	Phe	Phe	Gln	Thr 1430		Ser	Gly	Lys	Asn 1435		Ser	Val	Ala	Lys 144
	Glu Ser	Phe	Asn	Lys 1445		Val	Asn	Phe	Phe 1450		Gln	Lys	Pro	Glu 1455	
25	Leu His	s Asn	Phe 1460	Ser		Asn	Ser	Glu 1469		His	Ser	Asp	Ile 1470		Lys
23	Asn Lys	Met 147	Asp		Leu	Ser	Tyr 1480	Glu		Thr	qaA	Ile 1485	Val		His
	Lys Ile	e Leu		Glu	Ser	Val 149		Val	Gly	Thr	Gly 1500		Gln	Leu	Val
30	Thr Phe	Gln	Gly	Gln	Pro 1510	Glu		Asp	Glu	Lys 151		Lys	Glu	Pro	Thr 152
	Leu Le	ı Gly	Phe	His 1529	Thr		Ser	Gly	Lys 1530		Val	Lys	Ile	Ala 1535	
35	Glu Se	c Leu	Asp 1540	Lys		Lys	Asn	Leu 1549		Asp	Glu	Lys	Glu 1550		Gly
33	Thr Se	c Glu 155	Ile	Thr	Ser	Phe	Ser 1560	His		Trp	Ala	Lys 1569		Leu	Lys
	Tyr Arg	g Glu 70	Ala			157	Leu 5	Glu			1586)			
40	Ile Th:	r Ala	Ala	Pro	Lys 159		Lys	Glu	Met	Gln 159		Ser	Leu	Asn	Asn 160
	Asp Ly	s Asn	Leu	Val 1609	Ser		Glu	Thr	Val 1610		Pro	Pro	Lys	Leu 1615	
45	Ser As	p Asn	Leu 162		Arg	Gln	Thr	Glu 162		Leu	Lys	Thr	Ser 1630	Lys)	Ser
	Ile Ph	e Leu 163	Lys	Val	Гуs	Val	His 164	Glu		Val	Glu	Lys 164		Thr	Ala
	Lys Se 16	r Pro		Thr	Сув	Tyr	Thr		Gln	Ser	Pro 166		Ser	Val	Ile
50	Glu As		Ala	Leu	Ala 167	Phe		Thr	Ser	Cys 167		Arg	ГÀв	Thr	Ser 168
	Val Se	r Gln	Thr	Ser 168	Leu		Glu	Ala	Lys 169	Lys		Leu	Arg	Glu 1699	Gly
c c	Ile Ph	e Asp	Gly	Gln		Glu	Arg	Ile 170	Asn		Ala	Asp	Tyr 171	Val	
55	Asn Ty	r Leu 171	Tyr		Asn	Asn	Ser 172	Asn		Thr	Ile	Ala 172	Glu		qaA
	Lys As	n His		Ser	Glu	Lys 173	Gln		Thr	Tyr	Leu 174	Ser		Ser	Ser
60	Met Se	30 r Asn	Ser	Tyr		Tyr		Ser	Asp	Glu 175	Val		Asn	Asp	Ser 176
	1745 Gly Ty	r Leu	Ser	Lys	175 Asn		Leu	Asp	Ser			Glu	Pro	Val	

					1765					1770					1775	
	_			1780					1785					Val 1790		
5			1795	;				1800	ŀ				1805			
	Сув	Val 1810		Glu	Leu		Thr 1815		Ser	Ser	Pro	Cys 1820		Asn	Lys	Asn
10	Ala 1825	Ala	Ile	Lys		Ser 1830		Ser	Asn	Ser	Asn 1835		Phe	Glu	Val	Gly 184
	Pro	Pro			1845					1850)				1855	
				1860	Lys	Val			1865	;				Phe 1870		
15			1875	5				1880)				1885			
		1890	Ala	Gly			1895	5				1900)	Asp		
20	1909	Asn	Ser			1910)				1915	5		His		192
	Phe	Ala			1925	;				1930)			Gln	1935	i
				1940)				1945	5				Val 1950		
25			1955	5				1960)				1965			
		197	0				197	5				1980)	Ser		
30	198	5				1990)				1999	5		Gln		200
					2005	;				2010)			Val	2015	,
				2020	2				2025	5				Asn 2030	,	
35			203	5				204	0				204			
		205	Λ				205	5				206	0	Ser		
40	206	_				237	0				207	5		Gly		208
					208	5				209	0			Tyr	209	•
				210	0				210	5				Asp 2110)	
45			211	5				212	0				212			
		213	0				213	5				214	0	Ser		
50	214	5				215	0				215	5		Phe		216
					216	5				217	0			Val	217	5
				218	0				218	5				Val 219	U	
55			219	15				220	0				220			
		221	10				221	.5				222	0	Tyr		
60	222	Gli 25	ı Ala			223	30				223	5		Asp		224
	Thi	As _I	o Ser	Lys	Leu 224		Sez	His	: Ala	Thr 225	His	Ser	Leu	Phe	Thr 225	- Суғ 5

				2260)				226	5				Gly 227	0	
5	Arg	Gly	Glu 2275		Leu	Ile	Leu	Val 2280		Glu	Pro	Ser	11e 228	Lys 5	Arg	Asn
	Leu	Leu 229		Glu	Phe	Asp	Arg 2299		Ile	Glu	naA	Gln 2300		ГÀв	Ser	Leu
	Lys 2309		Ser	Lys	Ser	Thr 2310		Asp	Gly	Thr	Ile 231		Asp	Arg	Arg	Leu 232
10	Phe	Met	His	His	Val 2329		Leu	Glu	Pro	Ile 2330		Сув	Val	Pro	Phe 233	
	Thr	Thr	Lys	Glu 2340	_	Gln	Glu	Ile	Gln 234!		Pro	Asn	Phe	Thr 235		Pro
15	Gly	Gln	Glu 2355		Leu	Ser	ГÀЗ	Ser 2360		Leu	Tyr	Glu	His 236	Leu 5	Thr	Leu
	Glu	Lys 2370		Ser	Ser	Asn	Leu 2375		Val	Ser	Gly	His 2380		Phe	Tyr	Gln
	Val 2389		Ala	Thr	Arg	Asn 2390		Lys	Met	Arg	His 2399		Ile	Thr	Thr	Gly 240
20	Arg	Pro	Thr	Lys	Val 2409		Val	Pro	Pro	Phe 2410	_	Thr	Lys	Ser	His 2419	
	His	Arg	Val	Glu 2420		Cys	Val	Arg	Asn 2425		Asn	Leu	Glu	Glu 2430		Arg
25	Gln	Lys	Gln 2435		Ile	Asp	Gly	His 2440		Ser	qaA	Asp	Ser 244	Lys	Asn	Lys
	Ile	Asn 2450	_	Asn	Glu	Ile	His 2455		Phe	Asn	Lys	Asn 2460		Ser	Asn	Gln
	Ala 2465		Ala	Val	Thr	Phe 2470		Lys	Cys	Glu	Glu 2475		Pro	Leu	qaA	Leu 248
30	Ile	Thr	Ser	Leu	Gln 2485		Ala	Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2495	
	Lys	Lys	Gln	Arg 2500		Arg	Val	Phe	Pro 2505		Pro	Gly	Ser	Leu 2510		Leu
35	Ala	Lys	Thr 2515		Thr	Leu	Pro	Arg 2520		Ser	Leu	Lys	Ala 2525	Ala	Val	Gly
	Gly	Gln 2530		Pro	Ser		Cys 2535		His	Lys	Gln	Leu 2540		Thr	Tyr	Gly
	Val 2545		Lys	His	Сув	Ile 2550		Ile	Asn	Ser	Lys 2555		Ala	Glu	Ser	Phe 256
40	Gln	Phe	His	Thr	Glu 2569	_	Tyr	Phe	Gly	Lys 2570		Ser	Leu	Trp	Thr 2575	
	Lys	Gly	Ile	Gln 2580		Ala	Asp	Gly	Gly 2585		Leu	Ile	Pro	Ser 2590		Asp
45	_	_	2595	,	_			2600)				2605	Asp		
	Gly	Val 2610		Pro	Lys	Leu	1le 2615		Arg	Ile	Trp	Val 2620		naA	His	Tyr
	2625	5				2630)				2635	i		Phe		264
50					2645	5				2650)			Leu	2655	;
	Lys	Tyr	Arg	Tyr 2660		Thr	Ģlu	Ile	Asp 2665		Ser	Arg	Arg	Ser 2670		Ile
55	_	-	2675	5				2680)			_	2685			
		269	0				2699	5				2700)	Glu		
	270	5				2710)		•		2715	5		Ile		272
60					2725	5				2730)			Pro	2735	5
	Leu	Ala	Val	Leu	Lys	Asn	Gly	Arg	Leu	Thr	Val	Gly	Gln	Lys	Ile	Ile

				274					274	_				275		
	Leu	His	Gly 275!		Glu	Leu	Val	Gly 276		Pro	Asp	Ala	Cys 276		Pro	Let
5	Glu	Ala 277	Pro 0	Glu	Ser	Leu	Met 277		Lys	Ile	Ser	Ala 278		Ser	Thr	Arg
	Pro 278		Arg	Trp	Tyr	Thr 2790	-	Leu	Gly	Phe	Phe 279		Asp	Pro	Arg	Pro 280
10	Phe	Pro	Leu	Pro	Leu 2809		Ser	Leu	Phe	Ser 281		Gly	Gly	Asn	Val 281	_
	_		Asp	282	ס				282	5				2830	3	
	_		Ser 283	5				2840	0				284	5		
15	Lys	Glu 285	Ala O	Ala	Lys	туг	Val 285		Ala	Gln	Gln	Lys 2860		Leu	Glu	Ala
	Leu 286		Thr	Lys	Ile	Gln 2870		Glu	Phe	Glu	Glu 2879		Glu	Glu	Asn	Thr 288
20	Thr	Lys	Pro	Tyr	Leu 2889		Ser	Arg	Ala	Leu 2890		Arg	Gln	Gln	Val 289	_
			Gln	2900) _				290	5			-	2910)	
			Ala 2919	5				2920)				2925	5		
25		2930					2935	5				2940)			
	294	5	Glu		_	2950)				2955	5		_		296
30	_		Ser	_	2965	5				2970)		_		2975	5
	_		Lys	2980)				2989	5				2990)	
			Asp 2995	5	_			3000)				3005	5		
35	_	3010					3015	5				3020)	_		
	3025	5	Leu			3030)	_			3035	i				304
40			Glu		3045	i				3050)				3055	i
			Lys	3060)	_			3065	5			_	3070)	
45			Ile 3075	5				3080)		-	_	3085	,		
45		3090					3095	5				3100)			
	3105	5	Ile	_		3110)				3115	;				312
50			Ser		3125	5	_	-		3130	1	_		_	3135	;
			Phe	3140)				3145	i				3150		
==			Phe 3155	5				3160)				3165	5		
55		3170			_		3175	5			_	3180)			
	3185	5	Asn			3190)				3195	;				320
60			Tyr		3205	5				3210)	-		_	3215	i
	met	ser	Ser			Cys				_				Leu 3230		Leu

	Cys	Mec	323		AIG	тув	ser	3240		1111	PIO	Vai	324		GIII	Mec	
5	Thr	Ser 3250		Ser	Сув	Lys	Gly 325		Lys	Glu	Ile	Asp 326		Gln	Lys	Asn	
	Cys 3265		Lys	Arg	Arg	Ala 327		Asp	Phe	Leu	Ser 327		Leu	Pro	Leu	Pro 328	
	Pro	Pro	Val	Ser	Pro 3289		Сув	Thr	Phe	Val 3290		Pro	Ala	Ala	Gln 3295		
10	Ala	Phe	Gln	Pro 3300	Pro	-	Ser	Сув	Gly 3309	Thi		Tyr	Glu	Thr 3310	Pro		
	Lys	Lys	Lys 3319	Glu		Asn	Ser	Pro 3320	Gln		Thr	Pro	Phe 3325	Lys	Lys	Phe	
15	Asn	Glu 3330	Ile	-	Leu	Leu	Glu 3335	Ser		Ser	Ile	Ala 3340	Asp		Glu	Leu	
13		Leu		Asn	Thr		Ala		Leu	Ser	-	Ser		Gly	Glu	-	
	3345 Gln		Ile	Ser				ser	Thr	_			Pro	Thr	Ser		
20	Glu	Asp	Tyr				Lys	Arg	_	-		Thr	Ser		3375 Ile		
	Glu	Gln	Glu	3380 Ser		Gln	Ala	Ser	3385 Thr		Glu	Cys	Glu	3390 Lys) Asn	Lys	
	Gln	Asp	3395 Thr		Thr	Thr	Lys	3400 Lys		Ile			3405	5			
25		3410)				3415	5									
			(2)	INF	ORMA	TION	FOR	SEC	ID	NO:1	4:						
30		(i			ICE C												
30			(B)	TYPE	TH: : nu	clei	c ac	id									
					NDED LOGY			_									
35																	
					E/KE												
					ER I		ITAM	ON:	2F p	rime	r						
40		(x	i) S	EQUE	NCE	DESC	RIPT	: NOI	SEQ	ID	NO:1	4:					
	TGAG	TTTT	AC C	TCAG	TCAC	A											20
45			(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:1	6:						
45		(i		_	CE C												
					TH: : nu				B								
50			(C)	STRA	NDED LOGY	NESS	: si	ngle									
50			(1)	TOPO	LOGI	: 11	near										
		(x	i) S	EQUE	NCE	DESC	RIPT	ION:	SEQ	ID :	NO:1	6 :					
55	CAGG	AAAC.	AG C	TATG	ACCC	T GT	GACG	TACT	GGG	TTTT	TAG	С					41
			(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:1	7:						
60				_	CE C												
60					TH: : nu				8								
					NDED												

	(D) TOPOLOGY: linear	
5	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 3FII primer	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:	
	GATCTTTAAC TGTTCTGGGT CACA	24
	(2) INFORMATION FOR SEQ ID NO:18:	
15	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
20		
25	 (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 3RII primer (XI) SEQUENCE DESCRIPTION: SEQ ID NO:18: 	
		22
2.0	CCCAGCATGA CACAATTAAT GA	
30	(2) INFORMATION FOR SEQ ID NO:19:	
35	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 44 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
40	(A) NAME/KEY: (B) LOCATION:	
40	(D) OTHER INFORMATION: 4F/M 13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:	
45	TGTAAAACGA CGGCCAGTAG AATGCAAATT TATAATCCAG AGTA	44
	(2) INFORMATION FOR SEQ ID NO:20:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
55	(2)	
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 4R-1A primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	

163

ATCAGATTCA TCTTTATAGA AC

22

	(2) INFORMATION FOR SEQ ID NO:21:	
5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	
10	(D) TOPOLOGY: linear	
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 5+6F/M13F primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:	
	TGTAAAACGA CGGCCAGTTG TGTTGGCATT TTAAACATCA	40
20	(2) INFORMATION FOR SEQ ID NO:22:	
25	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 38 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
30	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 5+6R/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:	
35	CAGGAAACAG CTATGACCCA GGGCAAAGGT ATAACGCT	38
	(2) INFORMATION FOR SEQ ID NO:23:	
40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 38 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
45		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 7F/M13F primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:	
	TGTAAAACGA CGGCCAGTTA AGTGAAATAA AGAGTGAA	38
	(2) INFORMATION FOR SEQ ID NO:24:	
55	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 36 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
60	(C) STRANDEDNESS: SINGLE (D) TOPOLOGY: linear	

F	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 7R/M13R primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:	
	CAGGAAACAG CTATGACCAG AAGTATTAGA GATGAC	36
10	(2) INFORMATION FOR SEQ ID NO:25:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
20	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 8F/M13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:	
25	TGTAAAACGA CGGCCAGTGC CATATCTTAC CACCTTGTGA	40
	(2) INFORMATION FOR SEQ ID NO:26:	
30	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
35	(ix) FEATURE:	
4.0	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 8FIA primer	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:	
	TTGCATTCTA GTGATAATAT AC	22
45	(2) INFORMATION FOR SEQ ID NO:27:	
50	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 19 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
55	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 8RIA primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:	19
60	AATTGTTAGC AATTTCAAC (2) INFORMATION FOR SEQ ID NO:28:	
	/w/ **** ****	

5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 9F/M13F primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28: TGTAAAACGA CGGCCAGTTG GACCTAGGTT GATTGCAGAT	40
20	 (2) INFORMATION FOR SEQ ID NO:29: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
25	(D) TOPOLOGY: linear	
30	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 9R/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:	40
	CAGGAAACAG CTATGACCTA AACTGAGATC ACGGGTGACA	10
35	(2) INFORMATION FOR SEQ ID NO:30:	
40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
45	(A) NAME/KEY: (B) LOCATION: (D) OTHFR INFORMATION: 10AF primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:	
50	GAATAATATA AATTATATGG CTTA	24
J 0	(2) INFORMATION FOR SEQ ID NO:31:	
55	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 37 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
60	(A) NAME/KEY:	
	(B) LOCATION:	

	(D) OTHER INFORMATION: 10AR/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:	
5	CAGGAAACAG CTATGACCCC TAGTCTTGCT AGTTCTT	37
	(2) INFORMATION FOR SEQ ID NO:32:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 42 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
12		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 10BF/M13F primer	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:	
	TGTAAAACGA CGGCCAGTAR CTGAAGTGGA ACCAAATGAT AC	42
25	(2) INFORMATION FOR SEQ ID NO:33:	
30	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 44 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
35	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 10BR/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:	44
40	CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA	44
	(2) INFORMATION FOR SEQ ID NO:34:	
45	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
50	(ix) FEATURE:	
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 10CF/M13F primer	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:	
	TGTAAAACGA CGGCCAGTCA GCATCTTGAA TCTCATACAG	40
60	(2) INFORMATION FOR SEQ ID NO:35:	
	(i) SEQUENCE CHARACTERISTICS:	

5	(A) LENGTH: 19 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
10	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 10CRII primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:	
	AGACAGAGGT ACCTGAATC	19
15	(2) INFORMATION FOR SEQ ID NO:36:	
20	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
25	(A) NAME/KBY: (B) LOCATION:	
	(D) OTHER INFORMATION: 11AF-M13 primer	
2.0	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:	
30	TGTAAAACGA CGGCCAGTTG GTACTTTAAT TTTGTCACTT	40
	(2) INFORMATION FOR SEQ ID NO:37:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 37 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
40	,	
45	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11AR-M13 primer(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:	
	CAGGAAACAG CTATGACCTG CAGGCATGAC AGAGAAT	37
50	(2) INFORMATION FOR SEQ ID NO:38:	
55	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
60	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11BF primer	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:	
	AAGAAGCAAA ATGTAATAAG GA	22
5	(2) INFORMATION FOR SEQ ID NO:39:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
15	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11BR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:	
20	CATTTAAAGC ACATACATCT TG	22
	(2) INFORMATION FOR SEQ ID NO:40:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	
30	(D) TOPOLOGY: linear	
35	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11CF primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:	
	TCTAGAGGCA AAGAATCATA C	21
40	(2) INFORMATION FOR SEQ ID NO:41:	
45	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
50	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11CR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:	22
55	CAAGATTATT CCTTTCATTA GC	
60	(2) INFORMATION FOR SEQ ID NO:42: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	

	(D) TOPOLOGY: linear		
5	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11DF primer		
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:		
10	AACCAAAACA CAAATCTAAG AG		22
	(2) INFORMATION FOR SEQ ID NO:43:		
15	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 23 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single		
20	(D) TOPOLOGY: linear		
	(A) NAME/KEY:		
	(B) LOCATION:(D) OTHER INFORMATION: 11DR primer		
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:		
	GTCATTTTTA TATGCTGCTT TAC		23
30	(2) INFORMATION FOR SEQ ID NO:44:		
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 		
40	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11EF primer(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:		
45	GGTTTTATAT GGAGACACAG G		21
	(2) INFORMATION FOR SEQ ID NO:45:		
50	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 23 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 		
55	(A) NAME/KEY:		
	(B) LOCATION: (D) OTHER INFORMATION: 11ER primer		
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:		
	GTATTTACAA TTTCAACACA AGC	•	23

	(2) INFORMATION FOR SEQ ID NO:46:	
5	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 20 base pairs(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
10	(D) TOPOLOGY: linear	
10		
	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 11FF primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:	
		20
	ATCACAGTTT TGGAGGTAGC	20
20	(2) INFORMATION FOR SEQ ID NO:47:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 21 base pairs	
25	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
23	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
30	(B) LOCATION:(D) OTHER INFORMATION: 11FR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:	
35	CTGACTTCCT GATTCTTCTA A	21
	(2) INFORMATION FOR SEQ ID NO:48:	
	(i) SEQUENCE CHARACTERISTICS:	
40	(A) LENGTH: 21 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
45		
47	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 11GF primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:	
	CTCAGATGTT ATTTTCCAAG C .	2
	(2) INFORMATION FOR SEQ ID NO:49:	
55		
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 21 base pairs	
	(A) LENGTH: 21 base parts (B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
60	(D) TOPOLOGY: linear	

	(A) NAME/KEY:	
	(B) LOCATION:	
5	(D) OTHER INFORMATION: 11GR primer	
_	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:	
	CTGTTAAATA ACCAGAAGCA C	21
10	(2) INFORMATION FOR SEQ ID NO:50:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 18 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
20	(B) LOCATION:	
	(D) OTHER INFORMATION: 11HF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:	
25	AGGTAGACAG CAGCAAGC	18
	(2) INFORMATION FOR SEQ ID NO:51:	
	(i) SEQUENCE CHARACTERISTICS:	
30	(A) LENGTH: 22 base pairs	
	(B) TYPE: nucleic acid	
	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
	(b) Topologi. Timear	
35	(ix) FEATURE:	
	(A) NAME/KEY: None	
	(B) LOCATION:	
40	(D) OTHER INFORMATION: 11HR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:	
	GTAATATCAG TTGGCATTTA TT	22
45	(2) INFORMATION FOR SEQ ID NO:52:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 21 base pairs	
	(B) TYPE: nucleic acid	
50	(C) STRANDEDMESS: single (D) TOPOLOGY: linear	
	(D) TOPOLOGI. TIMEAL	
	(A) NAME/KBY:	
55	(B) LOCATION:	
	(D) OTHER INFORMATION: 111F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:	
60	TGCAGAGGTA CATCCAATAA G	21
	(2) INFORMATION FOR SEQ ID NO:53:	

5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(A) NAME/KBY:(B) LOCATION:(D) OTHER INFORMATION: 11IR primer	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:	
	GATCAGTAAA TAGCAAGTCC G	21
	(2) INFORMATION FOR SEQ ID NO:54:	
20	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 23 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
25	(D) TOPOLOGY: linear	
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11JF primer	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:	
	TACTGAAAAT GAAGATAACA AAT	23
35	(2) INFORMATION FOR SEQ ID NO:55:	
40	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
45	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: !!JR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:	
50	ATTTTGTTCT TTCTTATGTC AG	22
	(2) INFORMATION FOR SEQ ID NO:56:	
55	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 35 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
60	(A) NAME/KEY.	
	(A) NAME/KEY: (B) LOCATION:	

	(D) OTHER INFORMATION: 11KF-M13 primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:	
5	TGTAAAACGA CGGCCAGTCT ACTAAAACGG AGCAA	35
	(2) INFORMATION FOR SEQ ID NO:57:	
10	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 35 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
15	(b) TOPOLOGI: IIIIear	
20	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11KR-M13 primer	
- •	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:	
	CAGGAAACAG CTATGACCGT ATGAAAACCC AACAG	35
25	(2) INFORMATION FOR SEQ ID NO:58:	
30	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
35	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11LF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:	
40	CACAAAATAC TGAAAGAAAG TG	22
	(2) INFORMATION FOR SEQ ID NO:59:	
45	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 19 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
50		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11LR primer	
55	(x1) SEQUENCE DESCRIPTION: SEQ ID NO:59:	
	GGCACCACAG TCTCAATAG	19
60	(2) INFORMATION FOR SEQ ID NO:60:	
30	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs	

	(B) TYPE: nucleic acid	
	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
5	(b) forohoof. IImaal	
	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 11MF primer	
10	(2)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:	
	GCAAAGACCC TAAAGTACAG	20
	GCAAAGACCC TAAAGTACAG	
15	(2) INFORMATION FOR SEQ ID NO:61:	
	(1) analysis of the property of the control of the	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs	
	(B) TYPE: nucleic acid	
20	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
25	(B) LOCATION:	
	(D) OTHER INFORMATION: 11MR primer	
	(xi) SEOUENCE DESCRIPTION: SEO ID NO:61:	
	(X1) SEQUENCE DESCRIPTION: SEQ ID NO:01:	
30	CATCAAATAT TCCTTCTCTA AG	22
	(2) INFORMATION FOR SEQ ID NO:62:	
	(i) SEQUENCE CHARACTERISTICS:	
35	(A) LENGTH: 35 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
40		
	(A) NAME, XEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 11NF-M13 primer	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:	
	(Ma) 22(22) 22 22 22 22 22 22 22 22 22 22 22 22	
	TGTAAAACGA CGGCCAGTGA AAATTCAGCC TTAGC	35
	(2) INFOPMATION FOR SEQ ID NO:63:	
50	(2) INPOP WITON FOR DIN ID NO.03.	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 35 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
55	(C) STRANDEDNESS: STRIGTE (D) TOPOLOGY: linear	
	(D) TOPOLOGI. LILLOI	
	(A) NAME/KEY:	
60	(B) LOCATION: (D) OTHER INFORMATION: 11NR-M13 primer	
00	(D) OTHER THEORETION, TIME HED PLANET	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:	

	CAGGAAACAG CTATGACCAT CAGAATGGTA GGAAT	35
5	(2) INFORMATION FOR SEQ ID NO:64:	
10	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
15	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 110F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:	
20	GTACTATAGC TGAAAATGAC AA	22
	(2) INFORMATION FOR SEQ ID NO:65:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 20 base pair(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
30		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 110R primer	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:	
	ACCACTGGCT ATCCTAAATG	20
40	(2) INFORMATION FOR SEQ ID NO:66: (i) SEQUENCE CHARACTERISTICS:	
45	(A) LENGTH: 20 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
50	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11PF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:	
55	TGAAGATATT TGCGTTGAGG	20
	(2) INFORMATION FOR SEQ ID NO:67:	
60	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	

5	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11PR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:	
10	GTCAGCAAAA ACCTTATGTG	20
	(2) INFORMATION FOR SEQ ID NO:68:	
15	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 21 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
20		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11QF primer	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:	
	ACGAAAATTA TGGCAGGTTG T	21
30	(2) INFORMATION FOR SEQ ID NO:69:	
30	(i) SEQUENCE CHARACTERISTICS:	
35	(A) LENGTH: 21 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
4.0	(A) NAME/KEY: (B) LOCATION:	
40	(D) OTHER INFORMATION: 11QR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:	
45	CTTGTCTTGC GTTTTGTAAT G	21
	(2) INFORMATION FOR SEQ ID NO:70:	
50	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
55	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11RF primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:	
60	GCTTCATAAG TCAGTCTCAT	20

	(2) INFORMATION FOR SEQ ID NO:71:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11RR primer	
1.5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:	
	TCAAATTCCT CTAACACTCC	20
20	(2) INFORMATION FOR SEQ ID NO:72:	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 35 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
25	(D) TOPOLOGY: linear	
30	<pre>(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11SF-M13 primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:</pre>	
	TGTAAAACGA CGGCCAGTTA CAGCAAGTGG AAAGC	35
35	(2) INFORMATION FOR SEQ ID NO:73:	
40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 37 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
45	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11SR-M13 primer	
. .	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:	
50	CAGGAAACAG CTATGACCAA GTTTCAGTTT TACCAAT	37
	(2) INFORMATION FOR SEQ ID NO:74:	
55	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
60	(D) TOPOLOGY: linear	

178

(A) NAME/KEY:

	(B) LOCATION:(D) OTHER INFORMATION: 11TF primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:	
	GTTCTTCAGA AAATAATCAC TC	22
1.0	(2) INFORMATION FOR SEQ ID NO:75:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs	
15	(B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(A) NAME/KEY:	
20	(D) OTHER INFORMATION: 11TR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:	
25	TGTAAAAAGA GAATGTGTGG C	21
23	(2) INFORMATION FOR SEQ ID NO:76:	
30	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 39 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
35	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 11UF-M13 primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:	
40	TGTAAAACGA CGGCCAGTAC TTTTTCTGAT GTTCCTGTG	39
	(2) INFORMATION FOR SEQ ID NO:77:	
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 39 base pairs (B) TYPE: nucleic acid	
50	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
50	(A) NAME/KEY:	
55	(B) LOCATION: (D) OTHER INFORMATION: 11UR-M13 primer	
J 5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:	
	CAGGAAACAG CTATGACCTA AAAATAGTGA TTGGCAACA	39
60	(2) INFORMATION FOR SEQ ID NO:78:	
	(i) SEQUENCE CHARACTERISTICS:	

5	(A) LENGTH: 42 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
10	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 12F/M13F primer(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:	
15	TGTAAAACGA CGGCCAGTAG TGGTGTTTTA AAGTGGTCAA AA (2) INFORMATION FOR SEQ ID NO:79:	42
20	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
25	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 12R/M13R primer	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79: CAGGAAACAG CTATGACCGG ATCCACCTGA GGTCAGAATA (2) INFORMATION FOR SEQ ID NO:80:	40
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
40		
45	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 13-2F primer(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:	
	TAACATTTAA GCATCCGTTA C	21
50	(2) INFORMATION FOR SEQ ID NO:81:	
55	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 28 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
60	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 13-2R primer	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:	
_	AAACGAGACT TTTCTCATAC TGTATTAG	28
5	(2) INFORMATION FOR SEQ ID NO:82:	
10	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
15	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 14F primer	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:	
20	ACCATGTAGC AAATGAGGGT CT	22
	(2) INFORMATION FOR SEQ ID NO:83:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
30	(D) TOPOLOGY: linear	
35	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 14AR primer(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:	
	GCTTTTGTCT GTTTTCCTCC AA	22
40	(2) INFORMATION FOR SEQ ID NO:84:	
45	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
50	(A) NAME, KEY: (B) LOCATION: (D) OTHER INFORMATION: 15-2F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:	
55	CCAGGGGTTG TGCTTTTTAA A	21
	(2) INFORMATION FOR SEQ ID NO:85: (i) SEQUENCE CHARACTERISTICS:	
60	(A) LENGTH: 38 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	

(D) TOPOLOGY: linear

5	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 15FUT/M13-R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:	
10	CAGGAAACAG CTATGACCAC TCTGTCATAA AAGCCATC	38
	(2) INFORMATION FOR SEQ ID NO:86:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
20		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 16AF primer	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:	
	TTTGGTTTGT TATAATTGTT TTTA	24
30	(2) INFORMATION FOR SEQ ID NO:87:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
40	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 16AR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:	
45	CCAACTTTTT AGTTCGAGAG	20
	(2) INFORMATION FOR SEQ ID NO:88:	
50	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 19 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
55	(N) STANT / VEV.	
	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 17F primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:	
	TTCAGTATCA TCCTATGTG	19

	(2) INFORMATION FOR SEQ ID NO:89:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(b) Topologi. 11Mai	
15	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 17AR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:	
	AGAAACCTTA ACCCATACTG	20
20	(2) INFORMATION FOR SEQ ID NO:90:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 39 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
25	(D) TOPOLOGY: linear	
30	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 18FUT/M13-AF primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:	
	·	39
35	TGTAAAACGA CGGCCAGTGA ATTCTAGAGT CACACTTCC	
	(2) INFORMATION FOR SEQ ID NO:91:	
40	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 38 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
45		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 18R/M13R primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:	
	CAGGAAACAG CTATGACCTT TAACTGAATC AATGACTG	38
	(2) INFORMATION FOR SEQ ID NO:92:	
55	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 41 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
60	(D) TOPOLOGY: linear	

	(A) NAME/KEY:	
	(B) LOCATION:	
5	(D) OTHER INFORMATION: 19F/M13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:	
	TGTAAAACGA CGGCCAGTAA GTGAATATTT TTAAGGCAGT T	41
10	(2) INFORMATION FOR SEQ ID NO:93:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 39 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
20	(B) LOCATION:	
	(D) OTHER INFORMATION: 19FUT/M13-R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:	
25	CAGGAAACAG CTATGACCAA GAGACCGAAA CTCCATCTC	39
	(2) INFORMATION FOR SEQ ID NO:94:	
	(i) SEQUENCE CHARACTERISTICS:	
30	(A) LENGTH: 38 base pairs	
	(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
	(C) STRANDEDNESS: SINGLE (D) TOPOLOGY: linear	
	(2) 101020011 12::0-2	
35		
	(A) NAME/KBY:	
	(B) LOCATION:(D) OTHER INFORMATION: 20F/M13F primer	
	(D) OTHER INFORMATION: 2017 HIST PITMEL	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:	
	TGTAAAACGA CGGCCAGTCA CTGTGCCTGG CCTGATAC	38
45	(2) INFORMATION FOR SEQ ID NO:95:	
13	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 39 base pairs	
	(B) TYPE: nucleic acid	
5 0	(C) STRY DEDNESS: single (D) TOPOLOGY: linear	
50	(D) TOPOLOGI: Timear	
	(A) NAME/KEY:	
55	(B) LOCATION:(D) OTHER INFORMATION: 20R/M13R primer	
ÞΣ	(D) OIRER INFORMATION. ZOR/MICK PLIMEL	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:	
60	CAGGAAACAG CTATGACCAT GTTAAATTCA AAGTCTCTA	39
00	(2) INFORMATION FOR SEQ ID NO:96:	

5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 39 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 21F/M13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:	
15	TGTAAAACGA CGGCCAGTGG GTGTTTTATG CTTGGTTCT	39
	(2) INFORMATION FOR SEQ ID NO:97:	
20	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
25	(a) was the (MDV)	
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 21R/M13R primer	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:	
	CAGGAAACAG CTATGACCCA TTTCAACATA TTCCTTCCTG	40
٥.	(2) INFORMATION FOR SEQ ID NO:98:	
35	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 19 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
40	(D) TOPOLOGY: linear	
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 22F-1A primer	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:	
	AACCACACCC TTAAGATGA	19
50	(2) INFORMATION FOR SEQ ID NO:99:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs	
55	(A) LENGTH: 20 base parts (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
60	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 22R-1A primer	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:	
5	GCATTAGTAG TGGATTTTGC	20
	(2) INFORMATION FOR SEQ ID NO:100:	
10	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
15	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 23FII primer	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:	
	TCACTTCCAT TGCATC	16
	(2) INFORMATION FOR SEQ ID NO:101:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 17 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
30	(D) TOPOLOGY: linear	
35	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 23RII primer(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:	
		17
40	TGCCAACTGG TAGCTCC (2) INFORMATION FOR SEQ ID NO:102:	
45	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
50	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 24 2F primer	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:	•
	TACAGTTAGC AGCGACAAAA	20
	(2) INFORMATION FOR SEQ ID NO:103:	
60	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 38 base pairs(B) TYPE: nucleic acid	

	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	-
5	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 24R/M13R primer	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:	
	CAGGAAACAG CTATGACCAT TTGCCAACTG GTAGCTCC	38
	(2) INFORMATION FOR SEQ ID NO:104:	
15	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 20 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
20	(D) TOPOLOGY: linear	
25	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 25F-7/23 primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:	
2.0	GCTTTCGCCA AATTCAGCTA	20
30	(2) INFORMATION FOR SEQ ID NO:105:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
40	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 25R-7/23 primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:	
45	TACCAAAATG TGTGGTGATG	20
	(2) INFORMATION FOR SEQ IL NO:106:	
50	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
55		
	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 26-2F primer	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:	

	and the second s	PCT/US98/16905
	WO 99/09164	-
	AATCACTGAT ACTGGTTTTG	20
5	(2) INFORMATION FOR SEQ ID NO:107:	
5	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 20 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
10	(D) TOPOLOGY: linear	
	(A) NAME/KEY:	
15	(B) LOCATION:(D) OTHER INFORMATION: 26-2R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:	
20	TATACTTACA GGAGCCACAT	20
20	(2) INFORMATION FOR SEQ ID NO:108:	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 18 base pairs	
25	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
30	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 27AF-1A primer	
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:	
35	CTGTGTGTAA TATTTGCG	18
	(2) INFORMATION FOR SEQ ID NO:109:	
40	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 40 base pairs	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
4.5	(D) TOPOLOGY: linear	
45	(A) NAME/KEY:	
	(B) LOCATION: (D) OTHER INFORMATION: 27AR/M13R primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:	
	CAGGAAACAG CTATGACGGC AAGTTCTTCG TCAGCTATTG	40
55	(2) INFORMATION FOR SEQ ID NO:110:	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 40 base pairs (B) TYPE: nucleic acid	
60	(C) STRANDEDNESS: single(D) TOPOLOGY: linear	

5	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 27BF/M13F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:	40
10	TGTAAAACGA CGGCCAGTGA ATTCTCCTCA GATGACTCCA	40
	(2) INFORMATION FOR SEQ ID NO:111:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 38 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
20	(A) NAME/KEY:(B) LOCATION:(D) OTHER INFORMATION: 27BR/M13R primer	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:	
25	CAGGAAACAG CTATGACCTC TTTGCTCATT GTGCAACA	38

WE CLAIM:

5 1. A genomic DNA containing a BRCA2 gene,

wherein the first twelve nucleotides beginning exon 5 are 5'-

TCCTGTTGTTCT-3' as set forth in SEQ. ID. NO: 1,

wherein nucleotides numbers 5782-5790 are GTTTGTGTT as set forth in SEQ. ID. NO: 4, and

wherein the last 20 nucleotides ending exon 15 are 5'CTGCGTGTTCTCATAAACAG-3' as set forth in SEQ. ID. NO: 2 and the first 20
nucleotides beginning exon 16 are 5'-CTGTATACGTATGGCGTTTC-3' as set forth in SEQ. ID. NO: 3.

15 2. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 A 1593 A 20 2457 T 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

3. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

30 1093 A 1342 C 1593 A 2457 T 35 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

4. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 C 1593 A 2457 T 2908 G 3199 A 3624 A 4035 C 7470 A 9079 G.

5. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 C 1342 A 20 1593 A 2457 C 2908 G 3199 G 3624 G 25 4035 T 7470 G 9079 G.

6. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 C 1593 A 35 2457 T 2908 G 3199 A 3624 G 4035 T 40 7470 G 9079 G.

7. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 C 1342 C 1593 G 2457 C 50 2908 A 3199 G

3624 A 4035 T 7470 A 5 9079 A.

8. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

10 2024 C 4553 C 4815 G 5841 T 5972 C.

15

9. A DNA comprising a BRCA2 coding sequence, wherein nucleotide numbers 643-666 are

CTTAGTGAAAGTCCTGTTGTTCTA and

wherein nucleotides numbers 5782-5790 are GTTTGTGTT.

20

10. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

1093 A 1342 A 1593 A 2457 T 2908 G 3199 A 30 3624 A 4035 T 7470 A 9079 G.

11. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

1093 A 1342 C 40 1593 A 2457 T 2908 G 3199 A 3624 A 45 4035 T 7470 A 9079 G as set forth in SEQ. ID. NO: 4.

12. The DNA according to claim 9 wherein the coding sequence nucleutides are as follows:

PCT/US98/16905 WO 99/09164

```
1093 A
          1342 C
          1593 A
 5
          2457 T
          2908 G
          3199 A
          3624 A
          4035 C
10
          7470 A
          9079 G
     as set forth in SEQ. ID. NO: 6.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 15 as follows:

```
1093 C
          1342 A
          1593 A
20
          2457 C
          2908 G
          3199 G
          3624 G
          4035 T
25
          7470 G
          9079 G
     as set forth in SEQ. ID. NO: 8.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 30 14. as follows:

```
1093 A
          1342 C
          1593 A
35
          2457 T
          2908 G
          3199 A
          3624 G
          4035 T
40
          7470 G
          9079 G
     as set forth in SEQ. ID. NO: 10.
```

The DNA according to claim 9 wherein the coding sequence nucleotides are 15. 45 as follows:

1093 C 1342 C 1593 G 50 2457 C

2908 A 3199 G 3624 A 4035 T 5 7470 A 9079 A as set forth in SEQ. ID. NO: 12.

The DNA according to claim 9 wherein the coding sequence nucleotides are 16. 10 as follows:

2024 C 4553 C 4815 G 15 5841 T 5972 C.

A BRCA2 protein having the following amino acids at the following peptide 17. numbers: 20

289 asparagine histidine 372 894 valine 991 asparagine 25 1852 valine 1853 cysteine 1854 valine 2951 alanine

- as set forth in SEQ. ID. NO: 5. 30
 - The BRCA2 protein having the following amino acids at the following peptide 18. numbers:
- 289 asparagine 35 372 asparagine 599 serine valine 894 991 asparagine 2951 alanine. 40
 - The BRCA2 protein having the following amino acids at the following peptide 19. numbers:
- 289 histidine 45 372 histidine valine 894 asparatic acid 991 2951 alanine
- as set forth in SEQ. ID. NO: 9. 50

20. The BRCA2 protein having the following amino acids at the following peptide numbers:

5 289 histidine

372 asparagine

894 isoleucine

991 aspartic acid

2951 threonine

- as set forth in SEQ. ID. NO: 13.
 - 21. The BRCA2 protein according to claims 17-20 having the following amino acids at the following peptide numbers:

15 59? serine

30

1442 serine

1915 threonine.

- 22. A haplotype of BRCA2 coding sequence (BRCA2^{omi 1}) as set forth in SEQ. ID.
- NO: 4 or a sequence complementary thereto.
 - 23. A BRCA2 protein comprising an amino acid sequence derived from BRCA2^{orni}

 1 as set forth in SEQ. ID. NO: 5.
- 24. A haplotype of BRCA2 coding sequence (BRCA2^{omi 2}) as set forth in SEQ. ID.
 NO: 6 or a sequence complementary thereto.
 - 25. A BRCA2 protein comprising an amino acid sequence derived from BRCA2^{omi}
 ² as set forth in SEQ. ID. NO: 7.
 - 26. A haplotype of BRCA2 coding sequence (BRCA2^{omi 3}) as set forth in SEQ. ID. NO: 8 or a sequence complementary thereto.
- 27. A BRCA2 protein comprising an amino acid sequence derived from BRCA2^{oml}
 3 as set forth in SEQ. ID. NO: 9.

28. A haplotype of BRCA2 coding sequence (BRCA2^{oml 4}) as set forth in SEQ. ID. NO: 10 or a sequence complementary thereto.

- ⁴ as set forth in SEQ. ID. NO: 11.
 - 30. A haplotype of BRCA2 coding sequence (BRCA2^{omi 5}) as set forth in SEQ. ID. NO: 12 or a sequence complementary thereto.
 - 31. A BRCA2 protein comprising an amino acid sequence derived from BRCA2^{omi}
 ⁵ as set forth in SEQ, ID, NO: 13.
- 32. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
 - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2
 coding sequence;
 - (b) sequencing said amplified DNA fragment;

10

20

25

- (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
- (d) comparing the sequence of said amplified DNA fragment to a
 BRCA2^(orni) DNA sequence selecting from the group consisting of SEQ.
 ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID.
 NO: 12, and their respective complementary sequences;
- (e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:
 - (i) AAT and CAT at position 1093,
 - (ii) CAT and AAT at position 1342,
 - (iii) TCA and TCG at position 1593,

5

10

15

25

- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2^(omi) DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 33. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
 - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
 - (b) sequencing said amplified DNA fragment;
 - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
 - (d) comparing the sequence of said amplified DNA fragment to a BRCA2^(omi) DNA sequence selecting from the group consisting of SEQ.
 ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;

(e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:

- (i) AAT and CAT at position 1093,
- (ii) CAT and AAT at position 1342,
- (iii) TCA and TCG at position 1593,
- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2^(omi) DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence; wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:
- (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,
 - (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively,
 - (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,

25

5

10

15

20

at position 2457, CAT and CAC occur at frequencies from (iv) about 75-85%, and from about 15-25%, respectively, at position 2908, GTA and ATA occur at frequencies from (v) about 85-95%, and from about 5-15%, respectively, 5 at position 3199, AAC and GAC occur at frequencies (vi) from about 75-85%, and from about 15-25%, respectively. at position 36⁴, AAA and AAG occur at frequencies (vii) from about 75-85%, and from about 15-25%, 10 respectively, at position 4035, GTT and GTC occur at frequencies from (viii) about 85-95%, and from about 5-15%, respectively, at position 7470, TCA and TCG occur at frequencies from (ix) about 75-85%, and from about 15-25%, respectively, and 15 at position 9079, GCC and ACC occur at frequencies (x) from about 85-95%, and from about 5-15%, respectively. A method of detecting an increased genetic susceptibility to breast and 34. ovarian cancer in an individual resulting from the presence of a mutation in the 20 BRCA2 coding sequence, comprising: amplifying a DNA or a fragment thereof of an individual's BRCA2 (a) coding sequence;

coding sequence is sufficiently sequenced to determine whether a mutation is present;

if necessary, repeating steps (a) and (b) until said individual's BRCA2

sequencing said amplified DNA fragment;

(b)

(c)

25

(d) comparing the sequence of said amplified DNA fragment to a
 BRCA2^(omi) DNA sequence selected from the group consisting of SEQ.

ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;

- determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2^(omi) DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
 - (i) AAT and CAT at position 1093,

5

10

15

20

- (ii) CAT and AAT at position 1342,
- (iii) TCA and TCG at position 1593,
- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 35. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
 - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
 - (b) sequencing said amplified DNA fragment;

(c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;

- (d) comparing the sequence of said amplified DNA fragment to a
 BRCA2^(omi) DNA sequence selected from the group consisting of: SEQ.
 ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID.
 NO: 12, and their respective complementary sequences;
- (e) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2^(omi) DNA sequence selected from the group consisting of: SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
 - (i) AAT and CAT at position 1093,
 - (ii) CAT and AAT at position 1342,
 - (iii) TCA and TCG at position 1593,
 - (iv) CAT and CAC at position 2457,
 - (v) GTA and ATA at position 2908,
 - (vi) AAC and GAC at position 3199,
 - (vii) AAA and AAG at position 3624,
 - (viii) GTT and GTC at position 4035,
 - (ix) TCA and TCG at position 7470, and
 - (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence, wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:

20

5

10

PCT/US98/16905

WO 99/09164

at position 1093, AAT and CAT occur at frequencies from (i) about 75-85%, and from about 15-25%, respectively, at position 1342, CAT and AAT occur at frequencies from (ii) about 35-45%, and from about 55-65%, respectively, 5 at position 1593, TCA and TCG occur at frequencies from (iii) about 85-95%, and from about 5-15%, respectively, at position 2457, CAT and CAC occur at frequencies from (iv) about 75-85%, and from about 15-25%, respectively, at position 2908, GTA and ATA occur at frequencies from (v) 10 about 85-95%, and from about 5-15%, respectively, at position 3199, AAC and GAC occur at frequencies (iv) from about 75-85%, and from about 15-25%, respectively, at position 3624, AAA and AAG occur at frequencies (vii) 15 from about 75-85%, and from about 15-25%, respectively, at position 4035, GTT and GTC occur at frequencies from (viii) about 85-95%, and from about 5-15%, respectively, at position 7470, TCA and TCG occur at frequencies from (ix) 20 about 75-85%, and from about 15-25%, respectively, and at position 9079, GCC and ACC occur at frequencies (x)

36. A method according to any of the claims 32-35 wherein the said amplifying is performed by annealing at least one oligonucleotide primer to said DNA fragment and extending the oligonucleotide primer by an agent for polymerization.

from about 25-95%, and from about 5-15%, respectively.

37. A method according to claim 36 wherein said oligonucleotide primer is directly or indirectly labeled with a radioactive label, a fluorescent label, a bioluminescent label, a chemiluminescent label, a metal chelator, or an enzyme label.

5

38. A BRCA2 coding sequence according to claims 32, wherein the codon pairs occur at the following frequencies:

10

(i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,

10

- (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively,
- (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,

15

- (iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,
- (v) at position 2908, GTA and ATA occur at frequencies from about 85-95%, and from about 5-15%, respectively,

(vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,

20

(vii) at position 3624, AAA and AAG occur at frequencies from about 75-85%, and from about 15-25%, respectively,

- (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively,
- (ix) at position 7470, TCA and TCG occur at frequencies from about 75-85%, and from about 15-25%, respectively, and
- (x) at position 9079, GCC and ACC occur at frequencies from about 85-95%, and from about 5-15%, respectively.

39. An oligonucleotide primer capable of hybridizing to a sample of BRCA2 gene, or its respective complementary sequences selected from the group consisting of SEQ. ID. NO: 14, 19, 22, 23, 25, 26, 29-76, 83, 85-88, 90, 91, 97, 98, 101, and 104-107.

- 40. A chip array having "n" elements for performing allele specific sequencebased techniques comprising a solid phase chip and oligonucleotides having "n" different nucleotide sequences,
- wherein "n" is an interger greater than or equal to ten,

5

10

15

20

30

wherein said oligonucleotides are bound to said solid phase chip in a manner which permits said oligonucleotides to effectively hybridize to complementary oligonucleotides or polynucleotides,

wherein oligonucleotides having different nucleotide sequence are bound to said solid phase chip at different locations so that a particular location on said solid phase chip exclusively binds oligonucleotides having a specific nucleotide sequence, and

wherein at least ten oligonucleotides are capable of specifically hybridizing to the BRCA2 DNA having the sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 10, SEQ. ID. NO: 12 or their respective complementary sequences, at least one oligonucleotide being capable of specifically hybridizing at each of the nucleotide positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, 9079, or complementary thereto.

- 25 41. A method of performing gene therapy on a patient, comprising:
 - a) contacting cancer cells *in vivo* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their respective complementary sequences
 - b) allowing the vector to enter the cancer cells, and
 - c) measuring a reduction in tumor growth.
 - 42. The method according to claim 41 wherein said cancer cells have a mutation in the BRCA2 gene.

43. The method according to claim 41 wherein said patient has a mutation in the BRCA2 gene of non-cancer cells.

5

10

- 44. A method of performing gene therapy on a patient or a sample, comprising:
- a) contacting cells *in vivo* or *in vitro* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their respective complementary sequences, and
- b) allowing the vector to enter the cells,
 wherein said patient has a reduced susceptibility for developing a cancer
 associated with a mutation in the BRCA2 gene.
- 15 45. A method according to claim 44 wherein said cells include healthy breast, ovarian or pancreatic tissues.
 - 46. A method according to claim 44 wherein a patient has an inherited mutation in the BRCA2 gene.

20

- 47. A method of treating a patient suspected of having a tumor, comprising:
- a) administering to a patient an effective amount of BRCA2 tumor growth inhibitor having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof;
 - b) allowing the patient's cells to take up the protein, and
 - c) measuring a reduction in tumor growth.
- 48. The method according to claim 47 wherein said tumor is a breast cancer, an ovarian cancer or a pancreatic cancer.
 - 49. The method according to claim 47 wherein said patient has an inherited mutation in the BRCA2 gene.

50. A method of preventing the formation or growth of a tumor, comprising:

a) adminstering to a patient an effective amount of BRCA2 tumor growth inhibiting protein having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof; and

- b) allowing the patient cells to take up the protein.
- 51. The method according to claim 31 wherein the protein is administered parenternally, by buccal adsorption or inhalation.
 - 52. A cloning vector comprising:

- (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any fragments thereof; and
- (b) one or more suitable regulatory sequences to induce replication and/or integration in a host cell.
- 53. An expression vector comprising a DNA sequence as set forth in SEQ. ID.
 NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any
 fragments thereof operatively linked to one or more promoter sequences capable of directing expression of said sequence in a host cell.
 - 54. A host cell transformed with the vector according to claim 52 or 53.
- 25 55. A BRCA2 polypeptide which is selected from the group consisting of:

 (a) a fragment of BRCA2 protein sequence as set forth in SEQ. ID. NO: 5,

 SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO:13;

 (b) an amino acid sequence which is substantially homologous to the BRCA2 protein sequence as set forth in SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID.

 NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO: 13;
 - (c) a molecule which has similar function to the BRCA2 protein; and (d) a fusion protein of (a), (b), or (c).

56. An anti-BRCA2 antibody wherein a molecule according to claims 17-21, 23, 25, 27, 29, 31, or 55 is used as an immunogen.

- 5 57. A diagnostic reagent comprising a molecule selected from the group consisting of:
 - (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO:
 - 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their complementary sequences;
 - (b) a nucleic acid fragment of (a) comprising at least 10 nucleotide in length;
- 10 (c) a sequence which hybridizes to (a) or (b);
 - (d) a polypeptide according to claim 17-21, 23, 25, 27, 29, 31, or 55; and
 - (e) an antibody which specifically binds to the polypeptide of (d).
- 58. A pharmaceutical composition comprising a molecule according to any one of the claims 17-21, 23, 25, 27, 29, 31, 55 in a pharmaceutically acceptable carrier.
 - 59. A pharmaceutical composition comprising a molecule according claim 56 in a pharmaceutically acceptable carrier.
- 20 60. A pharmaceutical composition comprising a molecule according to claim 57 in a pharmaceutically acceptable carrier.

Figure 1A

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA <u>ATĞ</u>CCTATTGGATCCAAAGAGAGGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

gggatttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAG AAGCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATT ACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTT CAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCT GTAAAAGAATTAGATAAATTCAAATTAGACTTAGgtaagtaatgcaatatggtagactgggg

tcactgaattattgtactgtttcagGAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACA GTGÄAAACTAAÄATGGATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

agggatttgctttgtTTATTTTAGTCCTGTTGTTCTACAATGTACACATGTAACACCACAAA GAGATAAGTCAGgtatgattaaaaacaatgctttttattctt

ttaacaattitccccttttttacccccagTGGTATGTGGGAGTTTGTTTCATACACCAAAGTTTGTG **AAGgtaaatatt**

TCTGAAAGTCTAGGAGCTGAGGTGGATCCTGÄTATGTCTTGGTCAAGTTCTTTAGC TACACCACCCACCCTTAGTTCTACTGTGCTCATAGgtaataata

ttttatcttacagTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

taaactataatttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGAŤTŤATCGCTTCTGTGACAGACAGTGAAAACACAAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct

ttaatgtgcttctgttttatactttaacagGATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATA GCTGCAAAGACCACATTGGAAAGTCAATGCCAAATGTCCTAGAAGATGAAGTATAT GAAACAGTTGTAGATACCTCTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTA GAACAAAAATCTACAAAAAGTAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATG TTGTATCTGAAGTGGAACCAAATGATACTGATCCATTAGATTCAAATGTAGCAAATC

Figure 1B

AGAAGCCCTTTGAGAGTGGAAGTGACAAAATCTCCAAGGAAGTTGTACCGTCTTTG GCCTGTGAATGGTCTCAACTAACCCTTTCAGGTCTAAATGGAGCCCAGATGGAGAA AATACCCCTATTGCATATTTCTTCATGTGACCAAAATATTTCAGAAAAAGACCTATTA GACACAGAGAACAAAGAAAGAAAGATTTTCTTACTTCAGAGAATTCTTTGCCACGT ATTTCTAGCCTACCAAAATCAGAGAAGCCATTAAATGAGGAAACAGTGGTAAATAA GAGAGATGAAGAGCAGCATCTTGAATCTCATACAGACTGCATTCTTGCAGTAAAGC AGGCAATATCTGGAACTTCTCCAGTGGCTTCTTCATTTCAGGGTATCAAAAAGTCTA TATTCAGAATAAGAGAATCACCTAAAGAGACTTTCAATGCAAGTTTTTCAGGTCATA TGACTGATCCAAACTTTAAAAAAGAAACTGAAGCCTCTGAAAGTGGACTGGAAATA CATACTGTTTGCTCACAGAAGGAGGACTCCTTATGTCCAAATTTAATTGATAATGGA AGCTGGCCAGCCACCACACAGAATTCTGTAGCTTTGAAGAATGCAGGTTTAAT ATCCACTTTGAAAAAGAAAACAAATAAGTTTATTTATGCTATACATGATGAAACATCT TATAAAGGAAAAAAATACCGAAAGACCAAAAATCAGAACTAATTAACTGTTCAGCC CAGTTTGAAGCAAATGCTTTTGAAGCACCACTTACATTTGCAAATGCTGATTCAGgta cctctatat

ttigtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAATGATCCAAAAAGCAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAATCGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA PAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT

Figure 1C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGGCTTTTAT TCTGCTCATGGCACAAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAAACTGTAAGTGAAAAAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGGAAACACTCAGATTAAAGA^GATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACAACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAACATCAAAAAGTATCTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAAAAGCAGCAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAAACT:CTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAATGACAAAAATCATC TCTCCGAAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT ACCATTCTGATGAGGTATATAATGATTCAGGATATCTCTCAAAAAATAAACTTGATT CTGGTATTGAGCCAGTATTGAAGAATGTTGAAGATCAAAAAAACACTAGTTTTTCCA AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGC ATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTT: GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT. AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

Figure 1D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCC~AGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG aagtgttcatttttacctttcgtgttgccaatca

eaeacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATAATAGAAAATCAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Ggtaaaattagctttttattata

aatatgtaatataaaattatttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT Exon 13 TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

ccccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTČAAĞAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAA.T TTTTAAAACTAAATCaCATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAAACAGACAAAACATTGATGGACATGGCTCTGATGATAGTAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

ttittgctaagtatttattctttgatag.^:TTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĀAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC **AAGTTCCCTCTGCgtgtccccataaacaggtatgtgt**

tuttettutgtgtgtgtgtttattttgtgtagGTGTTCTCATAAACAGCTGTATACGTATGGCGTTTCTAA ACATTGCATAAAAATTAACAGCAAAAATGCAGAGTCTTTTCAGTTTCACACTGAAGA TTATTTTGGTAAGGAAAGTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTG GATGGCTCATACCCTCCAATGATGGAAAGGCTGGAAAAGAAGAAGAATTTTATAGgtactct atgcaaaaagattgtgtgttaacttttatg

Figure 1E

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGGTTTATAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt

ttttgttttcacttttagATATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

CAAACTTGGATTCTTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT **ACAGgtatgatgtattcttgaaactta**

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGC AĂTĞAĂGAGAGGAAĞĂAAAGGAAGCAGCAAAATATGTGGAGGCCCAACAAAAGA GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtzzzzit agttatatggtacacattgttatttc

agittagtgaattaataatccttttgttttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TAACAAGACAGCAAGTTCGTGCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagaggagtaagaggacatataatgag

tittattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAA TCACAGGCAAATGTTGAATGATAAGAAACAAGCTCAGATCCAGTTGGAAATTAGGA AGGCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAAC CGTGTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAGAAAAAGATTCAGgtaagtatgt aaatgctttgttttta

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAAACTCAGTATCAACAACTACC Ggtacaaacctttcattgtaattttt

Figure 1F

Exon 24

gaatttttgttttgttttctgtagGTTTCAGATGAAATTTTATTTCAGATTTACCAGCCACGGGAGC CCCTTCACTTCAGCAAATTTTTAGATCCAGACTTTCAGCCATCTTGTTCTGAGGTGG ACCTAATAGGATTTGTCGTTTCTGTTGTGAAAAAAACAGgtaatgcacaatatagttaatttttttat tgattcttttaaaaaaacattgtct

Exon 25

taacattctttttttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

Exon 26

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCAAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

Exon 27

tacgttttcattttttatcagATGTCTTCTCCTAATTGTGAGATATATTATCAAAGTCCTTTATCA CTTTGTATGGCCAAAAGGAAGTCTGTTTCCACACCTGTCTCAGCCCAGATGACTTC AAAGTCTTGTAAAGGGGAGAAAGAGATTGATGACCAAAAGAACTGCAAAAAGAGAA GAGCCTTGGATTTCTTGAGTAGACTGCCTTTACCTCCACCTGTTAGTCCCATTTGTA CATTTGTTTCTCCGGCTGCACAGAAGGCATTTCAGCCACCAAGGAGTTGTGGCAC CAAATACGAAACACCCATAAAGAAAAAAGAACTGAATTCTCCTCAGATGACTCCATT TAAAAAATTCAATGAAATTTCTCTTTTGGAAAGTAATTCAATAGCTGACGAAGAACTT GCATTGATAAATACCCAAGCTCTTTTGTCTGGTTCAACAGGAGAAAAACAATTTATA TCTGTCAGTGAATCCACTAGGACTGCTCCCACCAGTTCAGAAGATTATCTCAGACT GAAACGACGTTGTACTACATCTCTGATCAAAGAACAGGAGGAGTTCCCAGGCCAGTA CGGAAGAATGTGAGAAAAAAAAGCAGGACACAATTACAACTAAAAAATATATCTAA GCATTTGCAAAGGCGACAATAAATTATTGACGCTTAACCTTTCCAGTTTATAAGACT **GGA**

Figure 2A

Exon 2

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA <u>ATĞ</u>ČCTAŤŤGGAŤCČAAĂGAGAGGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

gggattttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAG AAGCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATT ACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTT CAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCT GTAAAAGAATTAGATAAATTCAAATTAGACTTAGgtaagtaatgcaatatggtagactgggg

tcactgaattattgtactgtttcagGAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACA GTGĂAAACTAAĂATGĞATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

agggatttgctttgttttattttagTCCTGTTGTTCTACAATGTACACATGTAACACCACAAAGAG ATAAGTCAGgtatgattaaaaacaatgctttttattctt

ttaacaattttcccctttttttacccccagTGGTATGTGGGGAGTTTGTTTCATACACCAAAGTTTGTG AAGqtaaatatt

taatgatcagggcattictataaaaaataaactattitcttcctcccagGGTCGTCAGACACCAAAACATATT TCTGAAAGTCTAGGAGCTGAGGTGGATCCTGATATGTCTTGGTCAAGTTCTTTAGC TACACCACCCACCCTTAGTTCTACTGTGCTCATAGgtastaata

tittatcttacagTCAGA AATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

tazactataztttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGATTTATCGCTTCTGTGACAGACAGTGAAAACACAAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct.

ttaatgtgcttctgttttatactttaacagGATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATA GCTGCAAAGACCACATTGGAAAGTCAATGCCAAATGTCCTAGAAGATGAAGTATAT GAAACAGTTGTAGATACCTCTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTA GAACAAAAATCTACAAAAAGTAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATG TTGTATCTGAAGTGGAACCAAATGATACTGATCCATTAGATTCAAATGTAGCAAATC

Figure 2B

AGAAGCCCTTTGAGAGTGGAAGTGACAAAATCTCCAAGGAAGTTGTACCGTCTTTG GCCTGTGAATGGTCTCAACTAACCCTTTCAGGTCTAAATGGAGCCCAGATGGAGAA AATACCCCTATTGCATATTTCTTCATGTGACCAAAATATTTCAGAAAAAGACCTATTA GACACAGAGAACAAAGAAAGAAAGAATTTTCTTACTTCAGAGAATTCTTTGCCACGT ATTTCTAGCCTACCAAAATCAGAGAAGCCATTAAATGAGGAAACAGTGGTAAATAA GAGAGATGAAGACAGCATCTTGAATCTCATACAGACTGCATTCTTGCAGTAAAGC AGGCAATATCTGGAACTTCTCCAGTGGCTTCTTCATTTCAGGGTATCAAAAAGTCTA TATTCAGAATAAGAGAATCACCTAAAGAGACTTTCAATGCAAGTTTTTCAGGTCATA TGACTGATCCAAACTTTAAAAAAGAAACTGAAGCCTCTGAAAGTGGACTGGAAATA CATACTGTTTGCTCACAGAAGGAGGACTCCTTATGTCCAAATTTAATTGATAATGGA AGCTGGCCAGCCACCACACAGAATTC, GTAGCTTTGAAGAATGCAGGTTTAAT ATCCACTTTGAAAAAGAAAACAAATAAGTTTATTTATGCTATACATGATGAAACATCT TATAAAGGAAAAAAATACCGAAAGACCAAAAATCAGAACTAATTAACTGTTCAGCC CAGTTTGAAGCAAATGCTTTTGAAGCACCACITACATTTGCAAATGCTGATTCAGGt acctctgtct

Exon 11

titgtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAAATGATCCAAAAAGCAAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAA I CGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA AAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT

Figure 2C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGCTTTTAT TCTGCTCATGGCACAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAAACTGTAAGTGAAAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGGAAACACTCAGATTAAAGAAGATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACAACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGAGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAAACATCAAAAAGTATCTTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAACAGCAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAATGACAAAAATCATC TCTCCGAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGCATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTTT GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT TAAAAGTAACGAACATTCAGACCAGCTCACAAGAGAAAAAAATACTGCTATACGTA CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

Figure 2D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCCAAGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG aagtottcatttttacctttcgtgttgccaatca

Exon 12

aaaacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATÃATAGAAAATČAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Gotaaaattagctttttatttata

Exon 13

aatatgtaatataaaataattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

Exon 14

coccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTCAAGAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT TTTTAAAACTAAATCACATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAAACAGACAAAGCAAAACATTGATGGACATGGCTCTGATGATAGTAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

Exon 15

tttttgctaagtatttattctttgatagATTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĂAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC **AAGTTCCCTCTGCGTGTTCTCATAAACAGgtatgtqt**

Exon 16

titttcttttttgtgtgtgtttattttgtgtagCTGTATACGTATGGCGTTTCTAAACATTGCATAAAAATTA ACAGCĂĂĂĂTGCĂĞAĞTCTTTTCAGTTTCACACTGAAGATTATTTTGGTAAGGAAA GTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTGGATGGCTCATACCCTCC ${\tt AATGATGGAAAGGCTGGAAAAGAAGAATTTTATAGgtactctatgcaaaaagattgtgtgttaactttt}$ ato

Figure 2E

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGGTTTATAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt

ttttgttttcactttttagATATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

taaatcaatatatttattaatttgtccagATTTCTGCTAACAGTACTCGGCCTGCTCGCTGGTATAC CAAACTTGGATTCTTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT ACAGgtatgatgtattcttgaaactta

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGCAĂŤĞĂĂĞAĞAĞAAĞĂAAAĞĞAAĞCAĞCAAAATATĞTĞĞAĞĞCCCAACAAAAĞA GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtaaaatt acttatatggtacacattgttatttc

agttagtgaattaataatccttttgttttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TĂACĂĂGACAGCAAGŤTCGTĞCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagagagtaagaggacatataatgag

Exon 22

ttttattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAATCACAGGCAAATGTTGAATGATAAGAAAUAAGCTCAGATCCAGTTGGAAATTAGGA AGGCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAACC GTGTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAGAAAAAGATTCAGgtaagtatgta aatqctttgttttta

Exon 23

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGAGATACAGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAAACTCAGTATCAACAACTACC Ggtacaaacctttcattgtaattttt

Figure 2F

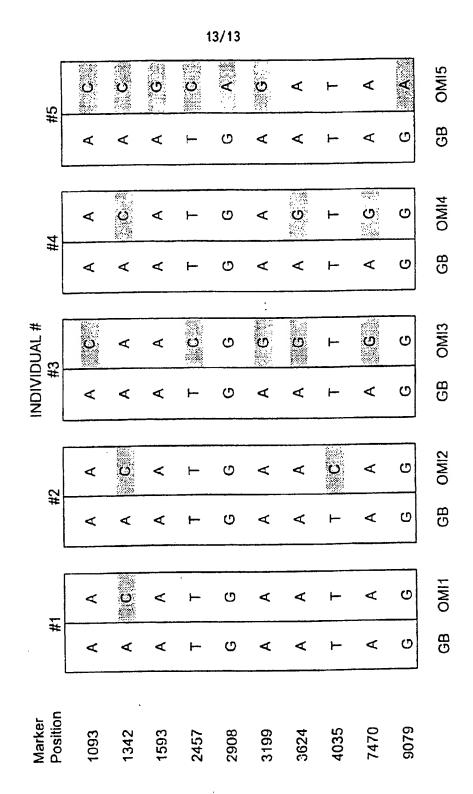
gaatttttgttttgttttctgtagGTTTCAGATGAAATTTTATTTCAGATTTACCAGCCACGGGAGC CCCTTCACTTCAGCAAATTTTTAGATCCAGACTTTCAGCCATCTTGTTCTGAGGTGG ACCTAATAGGATTTGTCGTTTCTGTTGAAAAAAACAGglaatgcacaatatagttaatttttttat tgattcttttaaaaaaacattgtct

taacattcttttcttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCAAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

tacgttttcatttttttatcagATGTCTTCTCCTAATTGTGAGATATATTATCAAAGTCCTTTATCA CTTTGTATGGCCAAAAGGAAGTCTGTTTCCACACCTGTCTCAGCCCAGATGACTTC AAAGTCTTGTAAAGGGGAGAAAGAGATTGATGACCAAAAGAACTGCAAAAAGAGAA GAGCCTTGGATTTCTTGAGTAGACTGCCTTTACCTCCACCTGTTAGTCCCATTTGTA CATTTGTTTCTCCGGCTGCACAGAAGGCATTTCAGCCACCAAGGAGTTGTGGCAC CAAATACGAAACACCCATAAAGAAAAAAGAACTGAATTCTCCTCAGATGACTCCATT TAAAAAATTCAATGAAATTTCTCTTTTGGAAAGTAATTCAATAGCTGACGAAGAACTT GCATTGATAAATACCCAAGCTCTTTTGTCTGGTTCAACAGGAGAAAAACAATTTATA TCTGTCAGTGAATCCACTAGGACTGCTCCCACCAGTTCAGAAGATTATCTCAGACT GAAACGACGTTGTACTACATCTCTGATCAAAGAACAGGAGAGTTCCCAGGCCAGTA CGGAAGAATGTGAGAAAAATAAGCAGGACACAATTACAACTAAAAAATATATCTAA GCATTTGCAAAGGCGACAATAAATTATTGACGCTTAACCTTTCCAGTTTATAAGACT **GGA**

FIGURE 3



inte onal Application No PCT/US 98/16905

A. CLASSIFICA			
Ä	TION OF SUBJECT MATTER 12N15/12 C07K14/47 C12Q1/68 61K38/17	A61K48/00	C07K16/18
According to Inte	mational Patent Classification (IPC) or to both national classificatio	n and IPC	
. FIELDS SEA	RCHED		
IPC 6 C	entation searched (classification system followed by classification of 12N CO7K C12Q A61K		U. C. I. de constant
	searched other than minimum documentation to the extent that suct		
Electronic data b	ase consulted during the International search (name of data base	and, where practical, search	Telline descrip
	S CONSIDERED TO BE RELEVANT		Solomonda elem No.
Category ° Cit	tation of document, with indication, where appropriate, of the relev	ant passages	Relevant to claim No.
Х	GB 2 307 477 A (CANCER RES CAMPAIG TECHNOLOGY LTD; DUKE UNIVERSITY) 28 May 1997	N	1,9, 55-60
	see figure 7, underlined sequences sheets 20/52, 27/52 and 31/52		1_16
X	TAVTIGIAN S V ET AL: "The complet gene and mutations in chromosome 13q-linked kindreds" NATURE GENETICS, vol. 3, no. 12, March 1996, pages XP002076942 cited in the application see table 1C		1-16, 32-39
	/	/	
X Further	documents are listed in the continuation of box C.	Y Patent family memb	ers are listed in annex.
"A" document considers "E" earlier document which is o citation o "O" document other me	defining the general state of the art which is not set to be of particular relevance current but published on or after the international set which may throw doubts on priority claim(s) or citied to establish the publication date of another or other special reason (as specified) to reterring to an oral disclosure, use, exhibition or lane.	or priority date and not it cited to understand the invention X* document of particular recannot be considered in involve an inventive stell Y* document of particular recannot be considered to comment to particular reconnections.	after the international filing date in conflict with the application but principle or theory underlying the devance; the claimed invention over or cannot be considered to p when the document is taken alone levence; the claimed invention or involve an inventive step when the with one or more other such document being obvious to a person skilled one arme patent family
	n the priority date claimed tual completion of the international search		ternational search report
	January 1999	02/02/1999	
Name and ma	hiling address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016	Authorized officer Cupido, M	

Inte onel Application No PCT/US 98/16905

(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	Relevant to claim No.	
egory °	Citation of document, with indication, where appropriate, of the relevant passages	Melevant to caunt No.	
	WO 97 22689 A (MYRIAD GENETICS INC ;UNIV PENNSYLVANIA (US); HSC RESEARCH & DEV LI) 26 June 1997 see SEQ ID NO:1 especially nucleotides 643-666 on page 90 and nucleotides 5782-5790 on page 98.	9,55-60	
, X	WO 97 30108 A (VANDERBILT UNIVERSITY; UNIVERSITY OF WASHINGTON) 21 August 1997 see SEQ ID NO:3 nucleotides 623-634, 5751-5759 and 7794-7834, for claim 9 see also nucleotides 616-634.	1,9	
	·		

....mational application No.

PCT/US 98/16905

Box I Observations where certain clair	ns were found unsearchable (Continuation of Item 1 of first sheet)
	stablished in respect of certain claims under Article 17(2)(a) for the following reasons:
Remark: Although claims	not required to be searched by this Authority, namely: 5 41-43 and 46-51 and partially claims 44 and 45 6 a method of treatment of the human/animal 6 ch has been carried out and based on the alleged 7 vector.
Claims Nos.: because they relate to parts of the Inte an extent that no meaningful Internation	ernational Application that do not comply with the prescribed requirements to such onal Search can be carried out, specifically:
	nd are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of in-	vention is lacking(Continuation of item 2 of first sheet)
This International Searching Authority found ma	ultiple inventions in this international application, as follows:
As all required additional search fees searchable claims.	were timely paid by the applicant, this International Search Report covers all
2. As all searchable claims could be see of any additional fee.	arched without effort justifying an additional fee, this Authority did not invitepayment
As only some of the required addition covers only those claims for which fe	nal search fees were timely paid by the applicant, this International Search Report ses were paid, specifically claims Nos.:
No required additional search fees we restricted to the invention first mention first mention.	. were timely paid by the applicant. Consequently, this International Search Report is consequently, this International Search Report is consed in the claims; it is covered by claims Nos.:
Remark on Protest	The additional search fees were accompanied by the applicant's protest. No protest accompanied the payment of additional search fees.

information on patent family members

Inter anal Application No PCT/US 98/16905

Patent document cited in search report		Publication date		atent family nember(s)	Publication date
GB 2307477	A	28-05-1997	AU CA EP WO	7635096 A 2238010 A 0858467 A 9719110 A	11-06-1997 29-05-1997 19-08-1998 29-05-1997
WO 9722689	Α	26-06-1997	US AU CA EP NO	5837492 A 1461597 A 2239733 A 0785216 A 982785 A	17-11-1998 14-07-1997 26-06-1997 23-07-1997 17-08-1998
WO 9730108	 A	21-08-1997	AU	2065397 A	02-09-199